Multidataset Study of Optimal Parameter and Uncertainty Estimation of a Land Surface Model with Bayesian Stochastic Inversion and Multicriteria Method

YOULONG XIA, MRINAL K. SEN, CHARLES S. JACKSON, AND PAUL L. STOFFA

Institute for Geophysics, The John A. and Katherine G. Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas

(Manuscript received 18 March 2003, in final form 5 May 2004)

ABSTRACT

This study evaluates the ability of Bayesian stochastic inversion (BSI) and multicriteria (MC) methods to search for the optimal parameter sets of the Chameleon Surface Model (CHASM) using prescribed forcing to simulate observed sensible and latent heat fluxes from seven measurement sites representative of six biomes including temperate coniferous forests, tropical forests, temperate and tropical grasslands, temperate crops, and semiarid grasslands. Calibration results with the BSI and MC show that estimated optimal values are very similar for the important parameters that are specific to the CHASM model. The model simulations based on estimated optimal parameter sets perform much better than the default parameter sets. Cross-validations for two tropical forest sites show that the calibrated parameters for one site can be transferred to another site within the same biome. The uncertainties of optimal parameters are obtained through BSI, which estimates a multidimensional posterior probability density function (PPD). Marginal PPD analyses show that nonoptimal choices of stomatal resistance would contribute most to model simulation errors at all sites, followed by ground and vegetation roughness length at six of seven sites. The impact of initial root-zone soil moisture and nonmosaic approach on estimation of optimal parameters and their uncertainties is discussed.

1. Introduction

Land surface modeling is considered to be one of the major causes of uncertainties in current climate change predictions (Houghton et al. 1996; Crossley et al. 2000). Moreover, results from the Project for Intercomparison of Land Surface Parameterization Schemes (PILPS) have revealed poor agreement among land surface schemes in representing key surface processes affecting water content and energy partitioning (Henderson-Sellers et al. 1996). Even when the same forcing is prescribed, there are still many factors that may cause this disparity. They include different model development philosophies (Henderson-Sellers 1996; Sellers et al. 1997), different model structures (Henderson-Sellers 1996), and different effective definitions of similar parameters (Desborough 1999). There have been continuing efforts to improve our understanding of model performance and how land surface schemes affect global climate simulations (Crossley et al. 2000; Desborough et al. 2001) and regional climate simulations (Zhang et al. 2001). As the structure and complexity of land surface models increase, the number of model parameters increase. Considerable research has been devoted to the development of automated methods for identifying optimal model parameter sets that reduce model uncertainties. Gupta et al. (1999) and Bastidas et al. (1999), for example, used a multicriteria (MC) calibration method to estimate acceptable optimal parameter sets for complex land surface schemes [e.g., the Biosphere–Atmosphere Transfer Scheme (BATS)]. The results showed that the BATS performed better when its parameter values were optimized using the MC method. Xia et al. (2002) used this method to investigate the relationship between model complexity and performance for the Cabauw dataset. Their results showed that complex models performed better than simple models when optimal model parameter values were used.

Estimates of parameters from a calibrated (optimized) complex land surface model are generally uncertain because 1) the observed data such as forcing data and surface energy flux data are uncertain, 2) the model never perfectly represents the land surface system and simplifications are needed in the parameterizations, and 3) model variables are ensemble averages while field measurements that are used for comparisons are ''instantaneous'' samples. Here we define uncertainty to be the choices of model parameter values that allow model simulations to exist within known observational or model errors. Therefore, the exercise of enumerating and evaluating the relative likelihood of different model parameter values is related to but is distinct in purpose from the ''global optimization'' techniques commonly

Corresponding author address: Youlong Xia, NOAA/Geophysical Fluid Dynamics Laboratory, and Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ 08542.
E-mail: youlong.xia@noaa.gov

used for calibration within the land surface modeling community such as the shuffled complex evolution method (SCE-UA; Duan et al. 1992, 1994), genetic algorithms (Wang 1991), or Bayesian recursive parameter estimation (Thiemann et al. 2001).

The application of statistical measures of uncertainty arising from multiple, nonlinearly related parameters typically uses a Monte Carlo Markov chain method based on the Metropolis–Hastings algorithm [Metropolis et al. 1953; Hastings 1970; for a modern treatment of this approach see Gelman et al. (2003)] and has been applied to surface hydrology by a number of authors (Kuczera and Parent 1998; Campbell et al. 1999; Bates and Campbell 2001) as well as models of the land surface (Franks and Beven 1997; Jackson et al. 2003). Other important examples of estimating land surface model parameter uncertainties include Alapaty et al. (1997) and Niyogi et al. (1998, 1999, 2002). The study by Franks and Beven (1997) used a Monte Carlo sampling of parameters within the Soil–Vegetation–Atmosphere Transfer scheme (SVAT) and generalized likelihood uncertainty estimation (GLUE) methodology to analyze uncertainties in land surface–atmosphere flux simulations for the First International Satellite Land Surface Climatology Project Field Experiment (FIFE) site and an Amazonian pasture site. The top 10% of 10 000 different parameter combinations were chosen to represent the uncertainty stemming from model parameters. The results showed that the range of model simulations of surface energy fluxes had typical widths of approximately one-third of the maximum observed fluxes for both sites. Franks and Beven (1997) reported that the short-term field campaigns represented by the datasets (6–21 August and 5–16 October 1987 at FIFE site; 16 October–2 November 1990 and 29 June–10 September 1991 at Amazon site) may be inadequate to specify parameter values characteristic of a site or area with precision.

Jackson et al. (2003) used a 1-yr dataset at the Cabauw site in the Netherlands and Bayesian stochastic inversion (BSI) to investigate the uncertainty of the Chameleon Surface Model (CHASM; Desborough 1999) for simulations of surface energy fluxes. Rather than specifying a given percentile of the top few percent of the parameter sets tested as was done within Franks and Beven (1997), a logic was introduced to select model parameter sets that were within uncertainties of the observed surface energy fluxes quantified by Beljaars and Bosveld (1997).

Both Xia et al. (2002) and Jackson et al. (2003) used only one site (Cabauw) as a case study. Therefore, multidatasets with different vegetation, soil, climate, and longer measurement periods are able to provide a more comprehensive comparison of the BSI and MC methods and quantification of parameter uncertainties within a complex land surface model. Here we use multidatasets from seven sites to compare the ability of the BSI and MC methods to search for the optimal parameter sets of the CHASM and use the BSI to analyze sources of parameter uncertainty for the CHASM. In addition, we discuss the uncertainties in specifying parameter values for the CHASM for the same biomes for different sites or for different periods for the same sites.

The following section gives a brief description of sites considered, the CHASM, and optimization methods. Section 3 describes the experiment design and analysis of parameter sensitivities. A comparison of the two optimization methods is shown in section 4, followed by an uncertainty analysis of the optimal parameters in section 5. The impacts of initial soil moisture and a nonmosaic approach on optimal parameters and their uncertainty estimation are given in section 6. The discussion and conclusions are presented in sections 7 and 8, respectively.

2. Sites, model, and optimization methods

Site characteristics such as site location, climate, vegetation type, soil type and measurement data characteristics such as data record length and observation intervals are described in section 2a, followed by a brief description of one-tile and two-tile CHASMs and basic parameterizations in section 2b. A brief description of Bayesian stochastic inversion and the multicriteria method is given in sections 2c and 2d, respectively.

a. Sites

The model forcing data and surface flux data used in this study were collected at seven sites. These sites were chosen based upon the data availability and for the different climate and vegetation characteristics. They represent midlatitude grasslands, midlatitude crops, tropical grasslands, tropical forests, midlatitude forests, and a semiarid shrubland. As suggested by Sen et al. (2001), these typical vegetations cover over 50% of the world's land area. At all sites, forcing data include downward longwave radiation (DLR), air temperature (T), relative humidity (RH), wind speed (V), precipitation (Precip) and incoming solar radiation (ISR) or net radiation (Rnet). The energy flux data include sensible and latent heat fluxes. Table 1 summarizes location, vegetation, climate, observed periods, and input data at seven sites.

1) ABRACOF TROPICAL FOREST SITE

The Abracof tropical forest data were taken at 10°5′S, 61°55′W, 80 km northeast of Ji-Parana in Rondonia, Brazil. The average tree height is 33 m and the soil is a red-yellow sandy loam (oxisol). The wet season is from December to April but the region experiences a dry season for several weeks during June and August when rainfall is less than 10 mm month^{-1}. Meteorological measurements were made on a 52-m-high tower. The forcing data were collected between June 1992 and December 1993 at 1-h intervals. Surface energy fluxes

TABLE 1. Description of seven sites and observed data (asterisk represents climate values).

Site name	Site location: lat, lon, elev (m)	Obs period	Obs intervals (min)	Vegetation type	Annual mean precipitation (mm)	Annual mean temperature (K)	Input data
Abracof	10°5′S 61°55′W 120 m	May 1992–1993	60	Tropical forest	1990	298	ISR, Rnet, *T*, RH, *V*, Precip
Abracop	10°45′S, 62°22′W 220 m	May 1992–Dec 1993	60	Tropical pasture	1985	297	ISR, Rnet, *T*, RH, *V*, Precip
Amazon	2°57′S 59°57′W 80 m	Jan 1997–Dec 1998	30	Tropical forest	1990	298	ISR, DLR, *T*, RH, *V*, Precip
ARMCART	36°36′N 97°29′W 318 m	Apr 1995–Aug 1995	30	Midlatitude crops	884*	284*	ISR, Rnet, *T*, RH, *V*, Precip
Cabauw	51°58′N 4°56′E −0.7 m	Jan 1987–Dec 1987	30	Midlatitude grassland	776	282	ISR, DLR, *T*, RH, *V*, Precip
Loobos	52°10′N 5°44′E 52 m	Jan 1997–Dec 1998	30	Midlatitude Scots pine	786	283	ISR, DLR, *T*, RH, *V*, Precip
Tucson	32°13′N 111°5′W 730 m	May 1993–May 1994	20	Semiarid grass and shrubs	275	293	ISR, DLR, *T*, RH, *V*, Precip

were also collected within two intensive observation periods between August and October in 1992 and between April and July in 1993 at 1-h intervals. Details can be found in Roberts et al. (1996) and online at http://www3.cptec.inpe.br/abracos.

2) ABRACOP TROPICAL GRASSLAND SITE

The Abracop tropical grassland site is located at 10°45′S, 62°22′W on a cattle ranch at an elevation of 220 m above sea level, about 50 km east-northeast of Ji-Parana. This site was deforested in 1977 and is at the center of a cleared area of about 50-km radius. The vegetation is grass and the soil is a red-yellow sandy loam latosol (oxisol). The area of bare soil is about 12%. The pasture had been burnt in the month prior to equipment installation, but it was not burnt again during the observation period. The forcing data were measured on a 6-m-high tower and quantities specifying the surface energy fluxes are the same as were collected at Abracof tropical forest site [see McWilliam et al. (1996) and http://www3.cptec.inpe.br/abracos].

3) AMAZON TROPICAL FOREST SITE

The Amazon tropical forest site is located at 2°57′S, 59°57′W in the Reserva Florestal Ducke, 25 km from Manaus, Brazil. The surrounding forest is undisturbed, and ground-level vegetation is below 1.2 m. Mean tree height is 35 m but some trees can be as high as 40 m, and the forest covers about 65%–70% of this area. The site was selected for its being representative of the natural vegetation and regional topography. The climatological average rainfall exhibits a marked seasonal dependence, with a mean monthly maximum about 350 mm in March, dropping to a mean monthly minimum about 100 mm in August. The forcing data were collected between January 1997 and December 1998 at 30-min intervals. The sensible and latent heat fluxes were collected during the same period at the same time intervals. A more extensive description of the site can be found in Shuttleworth (1984), Roberts et al. (1990), and online at http://www3.cptec.inpe.br/abracos.

4) ARMCART TEMPERATE CROP SITE

The U.S. Department of Energy's Atmospheric Radiation Measurement Program Clouds and Radiation Test Bed (ARMCART) crop site is located at 36°36′N, 97°29′W, near Lamont, Oklahoma. This site is surrounded by winter and summer crops. During the warm periods it suffers from lack of moisture. Meteorological data were measured on a 2.5-m-tall tower. The forcing data cover the 5-month period from 1 April to 25 August 1995, with a sampling interval of 30 min. At the same time, sensible and latent heat flux data were also collected [see Gupta et al. (1999) and http://www.arm.gov/docs/research/science/R00008.html].

5) CABAUW TEMPERATE GRASSLAND SITE

The data collected at Cabauw (51°58′N, 4°56′E, the Netherlands) are described in detail by Beljaars and Bosveld (1997). The data were measured on a 20-m-high tower surrounded by short grass divided by narrow ditches without significant obstacles or interruptions within a distance of 200 m from the tower. Beyond 200 m, some scattered houses and trees can be found. The climate in the area is characterized as moderate maritime with prevailing westerly wind. Annual precipitation is 776 mm, and annual mean air temperature is 282 K for the year 1987. The vegetation cover is close to 100% year-round. The soil contains 35%–55% clay. Forcing data and energy fluxes at 30-min intervals are collected for the entire year of 1987. The Cabauw data have been widely used to intercompare and investigate land surface models (Chen et al. 1997; Desborough 1999; Xia et al. 2002). The details of the data can also be found online at http://www.knmi.nl/onderzk/atmoond/cabauw/cabauw.html.

6) LOOBOS TEMPERATE CONIFEROUS FOREST SITE

The Loobos temperate coniferous forest site is located at 52°10′N, 5°44′E near Kootwijk in the Netherlands. The main tree species is Scots pine (*Pinus silvestris*). The forest extends to more than 1.5 km in all directions and is underlain by a grass understory growing in sand. The forcing data were collected at 30-min intervals on a 22-m-tall tower for both 1997 and 1998. The sensible and latent heat fluxes were collected with a 30-min interval for both years. More details can be found online at http://www.bgc-jena.mpg.de/public/carboeur/sites/index_s.html.

7) TUCSON SEMIARID SITE

The Tucson site is located at 32°13′N, 111°5′W in the semiarid, alluvial Sonoran Desert near Tucson, Arizona, on a gently sloping terrain (Unland et al. 1996). Annual precipitation is 275 mm at this site. The vegetation is very diverse and interspersed with patches of exposed rocky soil. The site has a fractional vegetation cover of 40%. Mean vegetation height is 1.2 m, ranging from low grass of a few centimeters to bushes of up to 7 m. The forcing data are similar to that at the ARM-CART site but at 20-min intervals, collected on a 10-m-tall tower from 1 June 1993 to 26 March 1994. Measurements of sensible and latent heat fluxes were made at the same frequency (see Unland et al. 1996).

b. Model

The CHASM (Desborough 1999; Pitman et al. 2003) land surface model has been used for offline intercomparison of the PILPS phase 2d (Schlosser et al. 2000; Slater et al. 2001) and phase 2e (Bowling et al. 2003) and simulations of the coupled global general circulation model (GCM; Desborough et al. 2001) and the regional climate model (Zhang et al. 2001). It was designed to explore the general aspects of land surface energy balance representation within a common modeling framework (Desborough 1999) that can be run in a variety of surface energy balance modes ranging from a complex mosaic-type structure (see Koster and Suarez 1992) all the way to the most simple zero energy balance formulation (Manabe 1969). Here we use the two-tile mosaic-type and one-tile nonmosaic-type representation. Within the mosaic-type representation the land–atmosphere interface is divided into two tiles. The first tile is a combination of bare ground and exposed snow with the second tile consisting of dense vegetation. The tiles may be of different sizes and the energy fluxes of each tile are area-weighted. Because separate surface balance is calculated for each tile, temperature variations may exist across the land–atmosphere interface. A prognostic bulk temperature for the storage of energy and a diagnostic skin temperature for the computation of surface energy fluxes are calculated for each tile. Snow fraction cover for both ground and foliage surfaces are calculated as functions of the snowpack depth, density, and the vegetation roughness length. The vegetation fraction is further divided into wet and dry fractions if canopy interception is considered. This model has explicit parameterizations for canopy resistance, canopy interception, vegetation transpiration, and bare-ground evaporation but has no explicit canopy-air space (see Pitman et al. 2003). Within the nonmosaic-type representation the land–atmosphere interface has only one tile.

CHASM uses the formulation of Manabe (1969) for the hydrologic component of the land surface in which the root zone is treated as a bucket with finite water-holding capacity. Any water accumulation beyond this capacity is assumed to be runoff. Except for moisture in the root zone, water can be stored as snow on the ground or on the canopy. Soil temperature is calculated within four soil layers using a finite-difference method and zero-flux boundary condition. Each tile has four evaporation sources for canopy evaporation, transpiration, bare-ground evaporation, and snow sublimation. Basic parameterizations related to latent heat fluxes are shown in the appendix.

c. Bayesian stochastic inversion

The BSI approach to quantifying uncertainties is based on the mathematics of conditional probabilities. That is, estimated parameter probabilities indicate the degree to which one set of model configurations perform better *relative* to other choices that have been considered. BSI controls which model configurations are tested. This is commonly done in a stochastic manner either through Monte Carlo (random) sampling or through other so-called importance sampling techniques such as

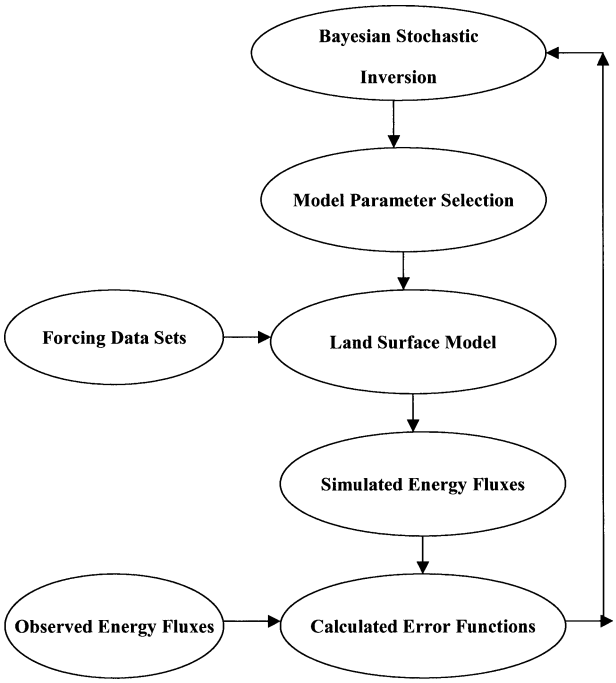

FIG. 1. A simple schematic diagram to use the BSI and CHASM land surface model.

very fast simulated annealing (VFSA) that was employed here. Every model configuration tested contributes to estimating a multidimensional probability distribution (called the posterior probability density function or PPD) by way of a likelihood function that provides a way to weigh different skill scores (defined in section 3a).

The PPD can be expressed mathematically as

$$\sigma(\mathbf{m}\,|\,\mathbf{d}_{\text{obs}}) = \frac{\exp[-sE(\mathbf{m})]p(\mathbf{m})}{\int \exp[-sE(\mathbf{m})]p(\mathbf{m})\,d\mathbf{m}}, \qquad (1)$$

where the vertical bar | is used in conditional probabilities to indicate that the probability for the left-side quantity is dependent on information provided by the limited sampling of quantities on the right side of the vertical bar. In this case Eq. (1) expressed the probability for specific combination of model parameter values **m** given the observations that are included in $\mathbf{d}_{\text{obs}}$. Here, $E(\mathbf{m})$ is the error function or skill score defined in section 3a, $\exp[-sE(\mathbf{m})]$ is the likelihood function with shaping factor s described later, $p(\mathbf{m})$ is the ''prior'' probability density function for **m** where one may specify in advance any constraints on choices of **m** outside the information provided by $\mathbf{d}_{\text{obs}}$. In our case we specify $p(\mathbf{m})$ as having a uniform distribution (or equal probability) for any realistic parameter choice. This prior provides a noninformative constraint for parameter values within the parameter search window. The denominator is a simple integration of the numerator so that the sum of all likelihood weights for all model combinations for a given parameter is unity.

TABLE 2. The calculated s values at seven measurement sites.

Site	Abracof	Abracop	Amazon	ARM-CART	Cabauw	Loobos	Tucson
S	5.7	7.3	6.5	6.8	16.1	15.4	23.1

Because the PPD is multidimensional, it is difficult to visualize. Therefore, a one-dimensional projection of the PPD (i.e., the marginal PPD) is usually displayed. Once the PPD is known, the parameter means or covariances can be obtained through multidimensional integrals of the general form

$$\mathbf{I} = \int f(\mathbf{m})\sigma(\mathbf{m}\,|\,\mathbf{d}_{\text{obs}})\,d\mathbf{m}, \qquad (2)$$

where $f(\mathbf{m}) = (\mathbf{m} - \langle\mathbf{m}\rangle)(\mathbf{m} - \langle\mathbf{m}\rangle)^{\text{T}}$ and $\langle\mathbf{m}\rangle$ is the vector of parameter means.

We use the VFSA algorithm to stochastically select parameter sets. VFSA is a form of importance sampling that helps to focus computational effort toward regions of parameter space that are the most significant portions of the PPD. As reported in Sen and Stoffa (1996), the VFSA algorithm convergence to the optimal parameter settings can be repeated a number of times within the BSI framework to approximate the multidimensional PPD, even when the relationship between parameters is nonlinear. Evaluations from all VFSA runs are used to estimate the PPD. A detailed description of the VFSA algorithm and its use can be found in Sen and Stoffa (1995, 1996) and Jackson et al. (2004), and a simple schematic diagram to use BSI and the CHASM is shown in Fig. 1.

Shaping factor s controls the relative weighting between different model configurations. The shaping factor is therefore related to uncertainty. Jackson et al. (2003) defined s in such a way as to force the PPD to reflect published estimates of the uncertainty in the observations ($s = 2/\Delta E_u$, where ΔE_u is the range in error function values associated with observational or model uncertainty). Because no such estimates exist for six of the seven PILPS datasets, here we use the root-mean-square errors (rmse) between observed data and simulations at the time scale of the observational data (~20–60 min) to define our best estimate of observational uncertainty (Table 6). Because this number also includes any systematic differences between model simulations and the observations, these estimates also include some notion of model uncertainty. If observational uncertainty is randomly distributed about the optimal selection, one may infer that the remaining systematic errors are most likely attributed to model biases. The calculated s values are shown in Table 2 for seven measurement sites. The s values vary from 5.7 to 23.1, depending on the site. Although s is related to the error between observed data and simulations, its choice is also affected

TABLE 3. Description and ranges of 12 CHASM parameters.

Parameter	Description	Min value	Max value
ALBG	Bare-ground albedo	0.05	0.40
ALBN	Snow albedo	0.50	1.00
ALBV	Vegetation albedo	0.05	0.40
LEFM	Max LAI	4.00	6.00
VEGM	Max fractional vegetation cover	0.70	1.00
VEGS	Fractional vegetation cover seasonality	0.23	0.26
RCMIN	Min canopy resistance (s m^{-1})	1.00	300
WRMAX	Available water-holding capacity (mm)	10.0	600
Z0G	Ground roughness length (m)	1.0×10^{-4}	0.01
Z0N	Snow roughness length (10^{-4} m)	1.0	6.0
Z0V	Vegetation roughness length (m) for grass	0.00	0.40
	Same for forest	0.80	2.50
TS	Initial soil temperature	275	310

by parameter sampling and therefore does not have a straightforward interpretation.

d. Multicriteria approach

The multicriteria (MC) parameter estimation methodology was developed by Gupta et al. (1998) from a single-criteria method of Duan et al. (1994) that is widely used in hydrological modeling. Gupta et al. (1998, 1999) describe a framework for the application of the multicriteria approach to the calibration of a physical-based model and present a case study in which this method is used to calibrate the BATS land surface model.

The basic idea of the MC method is to search for feasible parameter sets, in which parameter values simultaneously minimize multiple criteria according to one's objective function definition (Gupta et al. 1998). Since MC is used for a multiobjective problem (i.e., a vector of errors derived from many sources) it is unlikely to find a unique solution without stating how individual criteria should be weighted. Instead, there usually exist a range of solutions where moving from one solution to another results in improvement of one criterion while causing deterioration in another. This set is called the Pareto solution set which represents a range of the best solutions that can be found in the parameter space for each of the separate criteria. An efficient, population-based and optimized searching method, called the multiobjective complex evolution method, was developed by Yapo et al. (1998). It provides an approximate representation of the Pareto set in a single optimization run.

Prior to optimization, one needs to select feasible parameter ranges, objective function definition, and ''target'' observational fields. The MC algorithm will terminate when it has converged to the Pareto set.

3. Experimental design and analysis of parameter sensitivities

a. Experimental design

CHASM includes 12 soil and vegetation parameters that are available for calibration. Our initial set of experiments neglected to include the parameter for initial soil moisture. We have subsequently completed these experiments and remarked on its importance as well the influence of other choices pertaining to model structures that would be important to parameter optimization and uncertainty estimation in the discussion in section 7. The ranges and descriptions, and default values of these parameters are shown in Tables 3 and 4.

Observations of sensible and latent heat fluxes are used as the target observations for all seven sites. A ratio of variance of the errors (RVE) to the variance of

TABLE 4. Default values of 12 CHASM parameters.

Parameter	Abracof	Abracop	Amazon	ARMCART	Cabauw	Loobos	Tucson
ALBG	0.20	0.20	0.20	0.20	0.20	0.20	0.30
ALBN	0.75	0.75	0.75	0.75	0.75	0.75	0.75
ALBV	0.14	0.14	0.14	0.23	0.23	0.18	0.30
LEFM	6.00	4.00	6.00	4.00	4.00	4.00	4.00
VEGM	0.90	0.90	0.90	0.90	0.95	0.90	0.70
VEGS	0.25	0.25	0.25	0.25	0.25	0.25	0.25
RCMIN	50.0	50.0	50.0	40.0	40.0	40.0	40.0
WRMAX	234	234	234	141	141	200	122
Z0G	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Z0N	0.0004	0.004	0.0004	0.0004	0.0004	0.0004	0.0004
Z0V	2.00	0.20	2.00	0.15	0.15	2.00	0.02
TS	300	300	300	279	279	279	284

TABLE 5. Relative importance of CHASM parameters for seven sites (check marks indicate the important parameters).

Sites and parameters	Abracof	Abracop	Amazon	ARMCART	Cabauw	Loobos	Tucson
ALBG						√	√
ALBN							
ALBV	√	√	√		√	√	√
LEFM			√				
VEGM	√	√	√	√		√	√
VEGS							
RCMIN	√	√	√	√	√	√	√
WRMAX	√	√	√				√
Z0G	√	√	√	√	√		
Z0N							
Z0V	√	√	√	√	√		√
TS	√	√					

observations is used to define the mismatch between observations and model simulations for the MC and the BSI calibration. RVE is defined as

$$\mathrm{RVE} = \frac{\sum_{n=1}^{N} (\mathrm{obs}_n - \mathrm{sim}_n)^2}{\sum_{n=1}^{N} (\mathrm{obs}_n - \overline{\mathrm{obs}})^2}, \quad (3)$$

where N is number of observational data, obs_n is the observed data, sim_n is the simulation, and $\overline{\mathrm{obs}}$ is mean value of the observed data. MC and BSI are used for the seven sites to identify the best parameter sets that have minimum RVE. BSI is also used to estimate parameter uncertainty ranges.

b. Analysis of parameters

In order to understand what parameters are important for the CHASM at seven sites, we evaluated the marginal PPD. The most important parameters tended to have very sharp probability distributions. Those parameters that were marginally important may be more affected by statistical sampling issues. In these cases, the marginal PPD may be misleading. We therefore also considered linear sensitivity profiles (one factor at a time; Jackson et al. 2003), which do not have any sampling issues, to get a better perspective on the relative importance of individual parameters. The purpose of this analysis is to understand what parameters are sensitive or insensitive for the CHASM at seven sites. The experiment results are given in Table 5. The parameters that are identified as being important depend on the dominant physical processes that occur at each site. However, some parameters such as minimum stomatal resistance (RCMIN), vegetation albedo (ALBV), vegetation roughness length (Z0V), ground roughness length (Z0G), and vegetation cover fraction (VEGM) are important for almost all sites. RCMIN, Z0V, Z0G, and VEGM are closely related to the calculation of latent heat fluxes as shown in the appendix, while ALBV is closely related to the calculation of vegetation skin temperature, further influencing sensible and latent heat fluxes. In addition, ground albedo (ALBG), maximum water holding capacity (WRMAX), and initial ground temperature (TS) also have important influence for some sites. The other parameters, such as snow albedo (ALBN), maximum leaf area index (LEFM), vegetation fraction cover seasonality (VEGS), and snow roughness index (Z0N), show negligible effect on simulation errors for the CHASM for most of the sites. The parameters such as snow albedo and snow roughness length are not important because most of the sites have no snow (Abracof, Amazon, Abeacof, Tucson), have no snow during calibration period (ARMCART), or are not important for sensible and latent heat simulations (Cabauw, Loobos). Therefore, they are not sensitive parameters.

There are many methods to infer parameter sensitivity that have been applied in meteorological contexts, including the traditional perturbation (one parameter at a time) technique (e.g., Wilson et al. 1987; Bonan et al. 1993; Pitman 1994; Alapaty et al. 1997), variational (adjoint) methods (e.g., Skaggs and Barry 1996; Margulis and Entekhabi 2001), factorial methods (Henderson-Sellers 1993; Niyogi et al. 1999), Fourier amplitude sensitivity tests (e.g., Collins and Avissar 1994), muticriteria methods (e.g., Bastidas et al. 1996, 1999), reduced form model (Beringer et al. 2002), and response surface methods (Niyogi et al. 1998, 2002). Each of these techniques has benefits and drawbacks. For a thorough discussion of the pros and cons of the various methods, see Skaggs and Barry (1996) or Gao et al. (1996). Sensitive parameters are more dependent on specific land surface models, sites, and analysis methods used. However, that RCMIN, ALBV, and Z0V are important factors in this study is consistent with the previously cited studies.

4. Estimation of optimal parameter values for CHASM

a. Comparison of two methods

Table 6 shows optimal parameter sets derived by the BSI and MC methods at seven sites. The MC method

TABLE 6. Optimal parameter sets of the BSI and MC for 12 parameters of the CHASM land surface model (important parameters are boldfaced).

Sites and parameters	Abrafoc		Abracop		Amazon		ARMCART		Cabauw		Loobos		Tucson	
	BSI	MC	BSI	MC	BSI	MC	BSI	MC	BSI	MC	BSI	MC	BSI	MC
ALBG	0.39	0.30	0.39	0.34	0.40	0.36	0.40	0.25	0.26	0.17	**0.05**	**0.10**	**0.39**	**0.38**
ALBN	0.56	0.65	0.58	0.71	0.74	0.68	0.51	0.75	0.98	0.90	0.77	0.77	0.50	0.75
ALBV	**0.10**	**0.16**	**0.06**	**0.05**	**0.09**	**0.10**	0.05	0.15	**0.23**	**0.24**	**0.09**	**0.11**	**0.17**	**0.15**
LEFM	5.84	4.61	5.48	5.01	4.01	4.02	4.20	5.27	5.90	6.00	**5.95**	**5.97**	4.00	4.27
VEGM	**0.70**	**0.70**	**0.92**	**0.90**	**0.78**	**0.84**	0.78	0.86	**0.92**	**0.90**	**0.97**	**0.97**	**0.81**	**0.80**
VEGS	0.24	0.25	0.24	0.25	0.23	0.25	0.23	0.25	0.23	0.25	0.24	0.24	0.25	0.24
RCMIN	**20.0**	**19.3**	**33.6**	**35.3**	**50.1**	**45.9**	**17.8**	**22.7**	**14.2**	**12.6**	**43.8**	**43.1**	**96.8**	**99.7**
WRMAX	**599.3**	**589.9**	**593.8**	**595.7**	**567.6**	**525.5**	587.0	255.9	163.0	376.9	575.8	582.6	**269.3**	**262.8**
Z0G	1.0×10^{-4}	1.0×10^{-4}	1.3×10^{-4}	1.2×10^{-4}	1.0×10^{-4}	1.1×10^{-4}	1.0×10^{-4}	**1.0×10^{-4}**	**3.0×10^{-4}**	**3.2×10^{-4}**	9.0×10^{-3}	6.5×10^{-3}	1.2×10^{-3}	3.3×10^{-4}
Z0N	2.9×10^{-4}	2.9×10^{-4}	2.3×10^{-4}	2.5×10^{-4}	5.0×10^{-4}	3.5×10^{-4}	5.7×10^{-4}	2.6×10^{-4}	2.1×10^{-4}	3.0×10^{-4}	3.3×10^{-4}	5.3×10^{-4}	4.5×10^{-4}	3.0×10^{-4}
Z0V	**0.81**	**0.81**	**0.01**	**0.01**	**0.81**	**0.81**	**0.01**	**0.01**	**0.01**	**0.01**	0.83	1.62	**0.01**	**0.01**
TS	**275.2**	**278.1**	**276.0**	**277.2**	281.9	291.4	308.3	296.3	283.4	283.8	280.7	279.8	309.9	305.1

produces a number of parameter sets (called the Pareto set; Gupta et al. 1999) from which a preferred parameter set with minimum error values is chosen by a compromise method used by Leplastrier et al. (2002). In this compromise method we select from the Pareto set the parameter values that correspond to an equal weighting among the separate criteria (one for sensible heat flux and one for latent heat flux). A comparison of the optimal parameter sets obtained by BSI and MC methods shows that estimated optimal values are similar for the important parameters shown in Table 5. Because the important parameters largely determine simulations of sensible and latent heat fluxes, similar values for these parameters lead to similar simulations of sensible and latent heat fluxes (see Figs. 2, 3, and 4) at Abracof and Abracop (7–16 July 1993), Amazon (9–18 August 1998), ARMCART (4–13 July 1995), Cabauw (4–13 July 1987), Loobos (1–10 July 1998), and Tucson sites (13–22 August 1993). Here the CHASM was run for the entire length of the observational periods as shown in Table 1 and 10-day segments were selected for illustration purposes. Comparison of simulated energy fluxes with optimal parameter sets and default parameter sets taken from Desborough (1999) shows that at the tropical forest (Abracof and Amazon), tropical grassland (Abracop), and midlatitude grassland (Cabauw) and crop (ARMCART) sites, default parameter sets overestimate sensible heat fluxes and underestimate latent heat fluxes (see Fig. 4). These differences are due to the use of higher RCMIN and Z0V in the default parameter sets when compared with the optimal parameter sets. The large values for RCMIN and Z0V in the default parameter settings hinder the transpiration of moisture from vegetation and therefore favor sensible heat flux over latent heat flux to maintain the surface energy balance. At the midlatitude forest site (Loobos) and semiarid site (Tucson), the calibrated parameter sets and default parameter sets are similar, resulting in similar model simulations. Overall, the central conclusion is that both BSI and MC are effective methods for finding the optimal parameter sets that better reproduce the observations.

Root-mean-square errors and bias between simulated and observed data for whole simulation periods are shown in Table 7. Rmse varies between 12.8 and 49.8 W m^{-2} for sensible heat flux and between 22.9 and 70.0 W m^{-2} for latent heat flux, depending on the site and calibration method. When compared with default simulations, calibrated results have smaller rmse for all sites. At tropical forest and pasture sites (i.e., Abeacof, Abracop, Amazon) and midlatitude crops site (i.e., ARMCART), optimal parameter sets greatly reduce simulation biases for sensible and latent heat fluxes.

It should be noted that two problems exist in this study. The first one is that two methods produce significantly different values of WRMAX at ARMCART and Cabauw, significantly different values of Z0V at Loobos, significantly different values of ALBG at

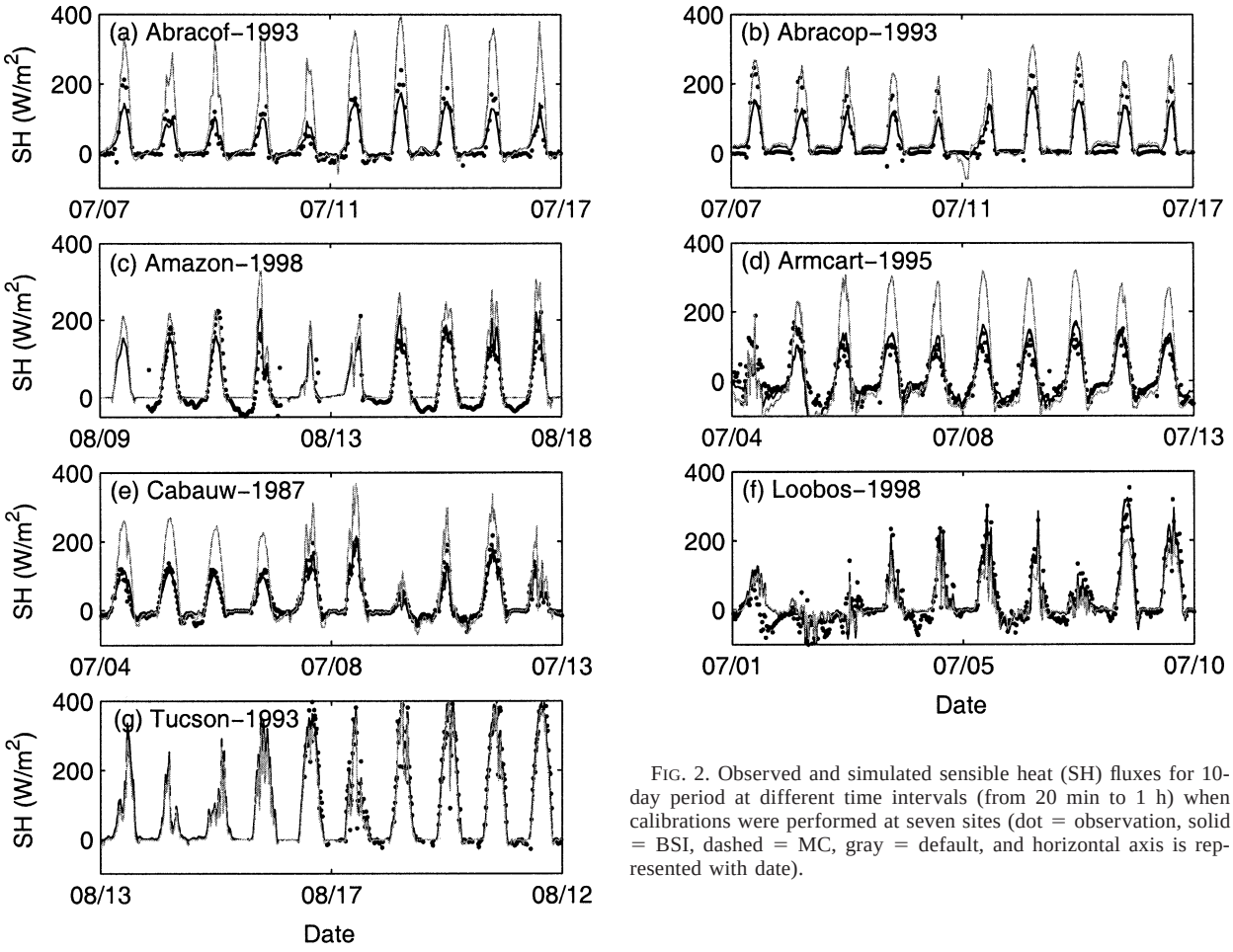

FIG. 2. Observed and simulated sensible heat (SH) fluxes for 10-day period at different time intervals (from 20 min to 1 h) when calibrations were performed at seven sites (dot = observation, solid = BSI, dashed = MC, gray = default, and horizontal axis is represented with date).

ARMCART, and significantly different values of TS at Amazon and ARMCART (Table 6). However, these parameters do not strongly affect offline sensible and latent heat flux simulations for the CHASM. Caution may be warranted before using calibrated parameter values for the parameters that are not strongly constrained by observations of surface energy fluxes. Feedbacks in coupled models may produce significant errors. In this case, we suggest using default values to replace the calibrated values for those parameters which are not strongly constrained. However, we still suggest using the calibrated values for the parameters that are strongly constrained by observations. The second problem is that some optimal parameter values generated in this study are not reasonable. For example, the value of optimal vegetation fraction (0.8) for Tucson is much larger than the default value and what one would assume for a semiarid region and that 0.7 for Abracof is much smaller than the default value and what one would assume to be the value for a tropical forest. The reason may be associated with the unreasonable model structure and with model structure–related interaction among parameters. Use of small parameter range (using local range instead of global range) in prior can avoid these unreasonable parameter values.

b. Spatial transferability of optimal parameters

Spatial transferability of the calibrated parameters for the same vegetation cover is an important and practical issue. Sen et al. (2001) have used calibrated parameters within the National Center for Atmospheric Research (NCAR) Community Climate Model, version 3 (CCM3), assuming calibrated parameters for selected locations could be used for the similar biomes at other locations. Two tropical forest sites (Abracof, Amazon) provide us with an opportunity to examine transferability of calibrated parameters for the same biomes. To assess the spatial transferability of the calibrated parameters, we cross-validate the parameters calibrated at Abracof using BSI and MC to calculate sensible and latent heat fluxes at the Amazon site and vice versa. Figure 5 shows the model simulations of sensible and latent heat fluxes over a 10-day period. The results demonstrate that sensible and latent heat fluxes are close to observed values at two sites and are better than results obtained using default parameter values. This evaluation exercise suggests that the calibrated parameters can be transferred from one site to another of the same biome type. Notwithstanding the potential importance of at-

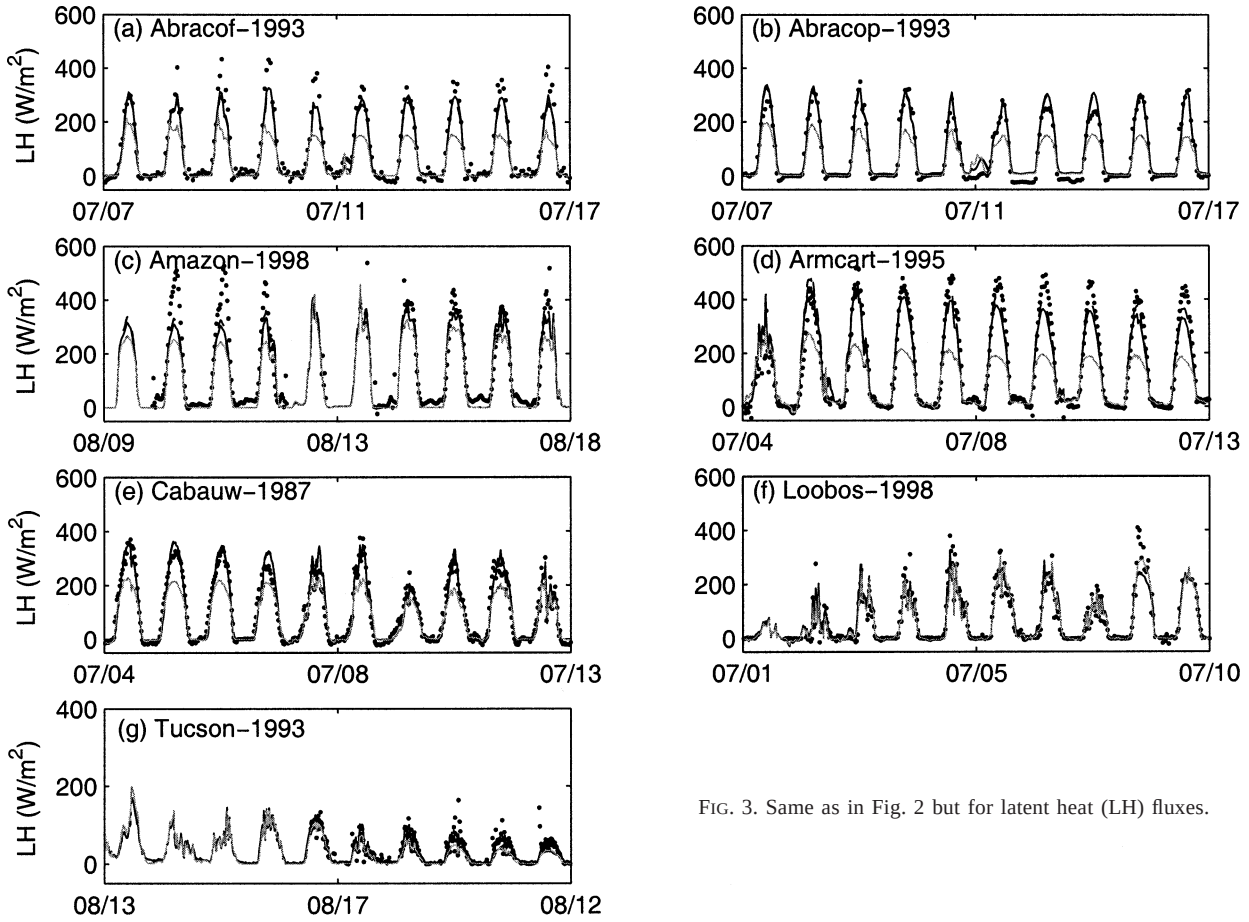

FIG. 3. Same as in Fig. 2 but for latent heat (LH) fluxes.

mospheric feedbacks, this exercise partly provides support for the use of offline calibrated parameter sets for whole vegetation classes within general circulation models (GCMs; Sen et al. 2001). If field observations can be made for 18 vegetation classes (from bare soil to forests), this application will improve GCM simulations and reduce climate model simulation uncertainties. However, the cross-validation exercise conducted here uses two datasets obtained within a few tens of kilometers of each other. Therefore, similar cross-validations may be required for testing whether or not results based on datasets within the same biome on different continents are interchangeable.

It is also still debatable whether the parameters calibrated for a given field site would be appropriate for a land surface model as applied to whole grid cells of a GCM. In addition, calibrated soil parameters may not be applicable to GCM simulations because soil types and characteristics are heterogeneous (Sen et al. 2001). Perhaps in cases where it is determined that specific soil parameters do not have a critical impact on GCM or land surface simulations, a typical value can be used.

TABLE 7. Rmse and bias between simulated and observed sensible (SH) and latent heat flux (LH) for seven measurement sites (W m^{-2}).

Method	BSI				MC			
Variable	SH		LH		SH		LH	
Site	Rmse	Bias	Rmse	Bias	Rmse	Bias	Rmse	Bias
Abracof	26.3	3.6	46.3	−20.0	26.3	3.1	49.2	−22.4
Abracop	29.5	0.3	55.4	−2.7	29.5	1.3	55.4	−3.5
Amazon	38.9	1.5	55.0	−11.0	39.3	−1.9	55.2	−10.2
ARMCART	49.8	7.3	70.0	−10.1	49.6	6.7	68.9	−9.8
Cabauw	13.1	2.4	23.1	1.9	12.8	2.7	22.9	2.0
Loobos	44.5	19.3	23.4	−4.1	44.0	16.6	24.0	−4.0
Tucson	34.1	6.5	27.4	−3.7	33.9	7.4	27.4	−3.5

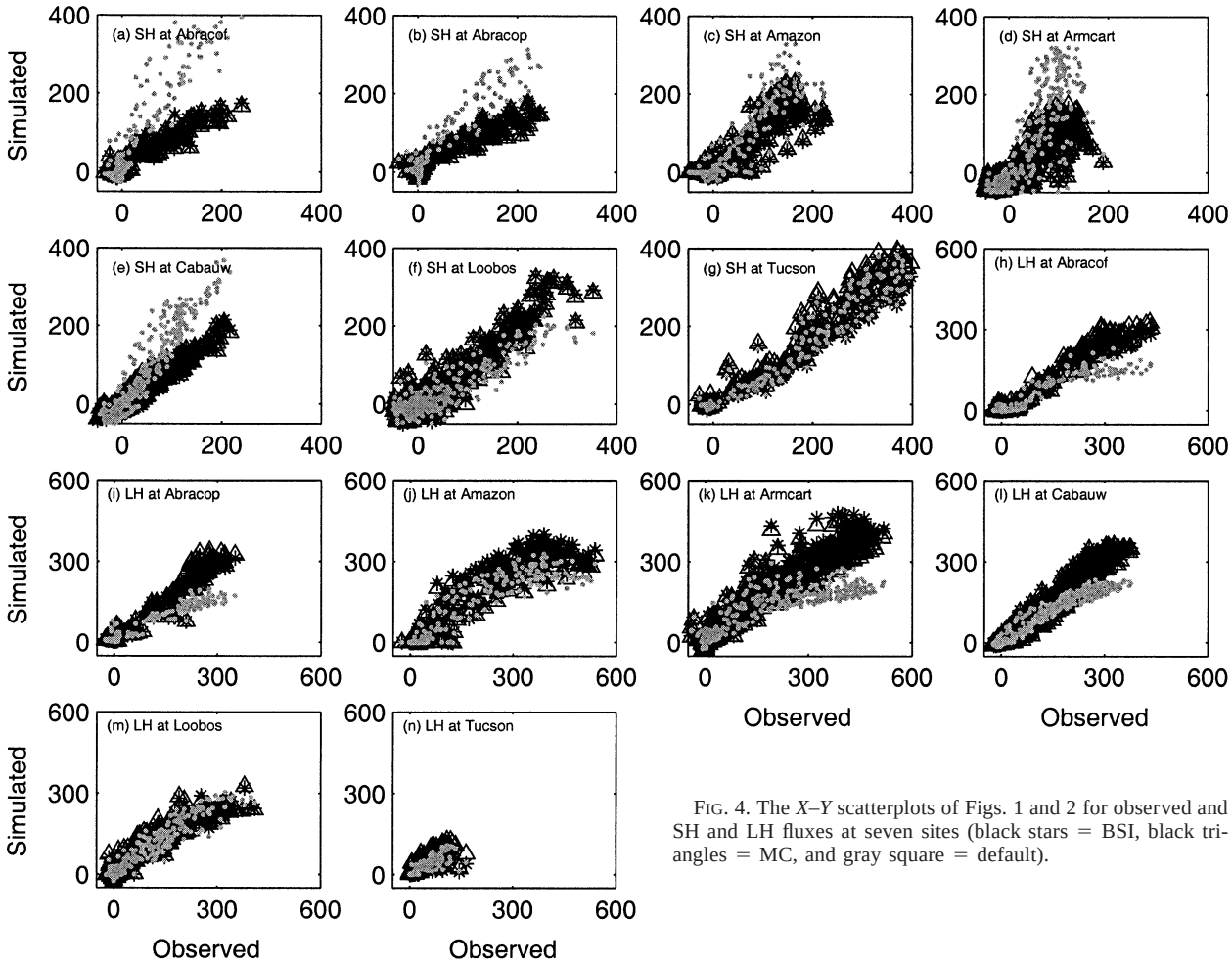

FIG. 4. The *X*–*Y* scatterplots of Figs. 1 and 2 for observed and SH and LH fluxes at seven sites (black stars = BSI, black triangles = MC, and gray square = default).

However, this will probably not be true for sites such as Abracof, Abracop, or Tucson where soil parameters show significant effect on simulations (i.e., maximum soil moisture-holding capacity).

5. Uncertainty analysis of optimal parameters

a. Marginal PPD

In Figures 6 and 7, we compare the marginal PPDs and parameter ranges at the 90% and 95% confidence levels for four parameters that are consistently important at the seven sites considered. Here, we only concentrated on discussing the results on the 95% confidence level. The results for grasslands at the Abracop, ARMCART, Cabauw, and Tucson sites are shown from the first row to the fourth row in Fig. 6, respectively. The results for forested areas Abracof, Amazon, and Loobos are shown from the first row to the third row in Fig. 7, respectively. Lines between two circles represent ranges of parameters at the 90% confidence level. Lines between two dots represent ranges of parameters at the 95% confidence level. The diamonds represent the best parameter sets for different parameters. Comparison of the marginal PPDs for these four parameters shows significant variations among midlatitude grasslands, tropical forests, tropical grasslands, and semiarid areas, associating different levels of uncertainty to specific parameter choices for different vegetation covers and climates. For grasslands and bushes (see Fig. 6) the marginal PPD of vegetation roughness length (e.g., Z0V) has maximum values (indicating maximum likelihood or possibility) at small values (e.g., $0–10^{-2}$ m). These small values are consistent with the optimal parameter values (repre-

TABLE 7. (*Extended*)

Default			
SH		LH	
Rmse	Bias	Rmse	Bias
75.4	35.2	82.7	−37.6
48.5	18.2	72.9	−21.9
45.7	9.0	62.2	−16.2
73.9	−6.4	93.2	−24.5
30.4	3.5	32.9	0.0
51.3	8.6	25.6	−4.2
45.3	−11.2	33.7	−1.3

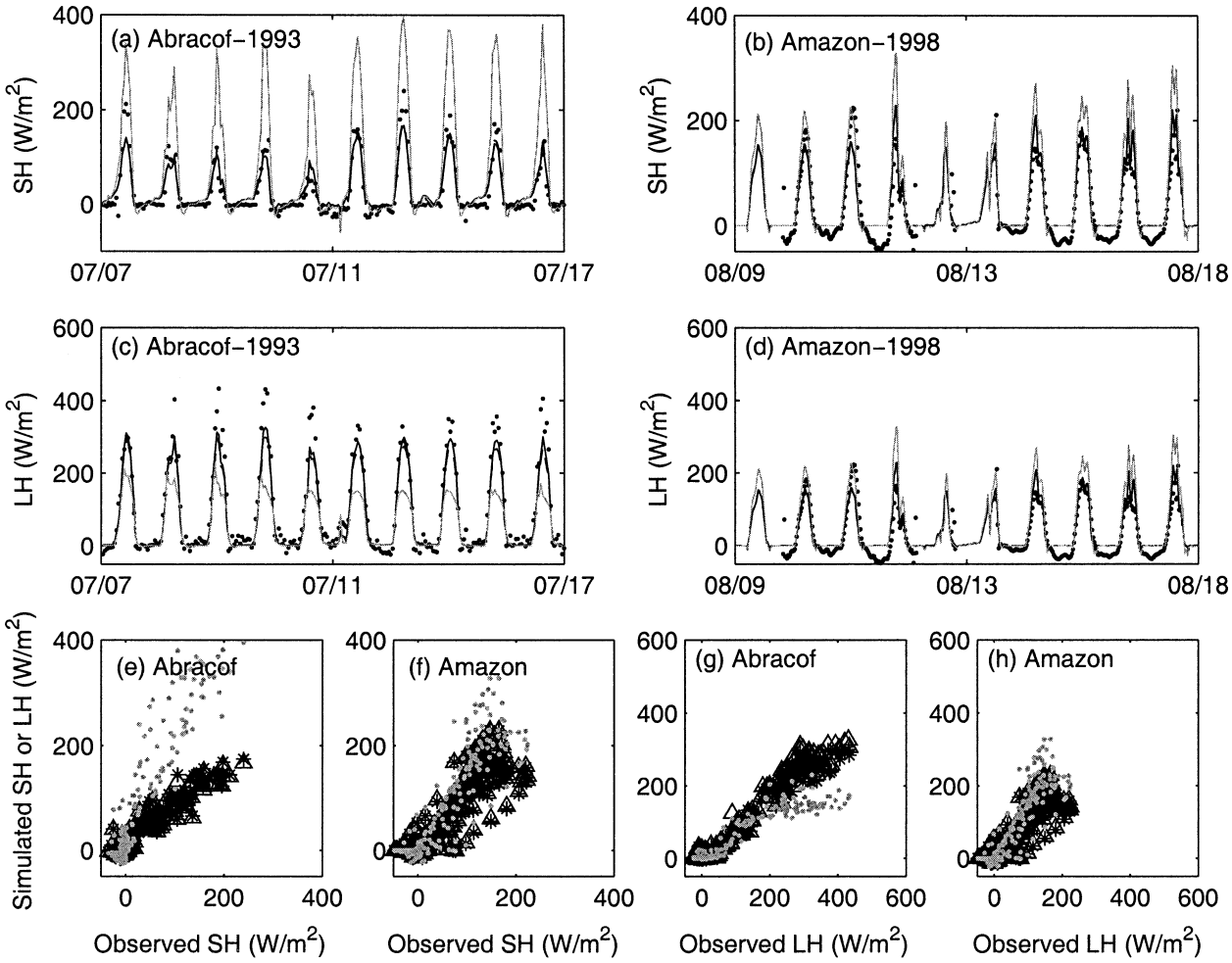

FIG. 5. Cross-validation showing observed and SH and LH fluxes for a 10-day period with different time intervals (from 20 min to 1 h) at two sites (dot = observation, solid = BSI, dashed = MC, gray = default, and horizontal axis indicates date). The optimal parameter sets at the Abracof site were used at the Amazon site and vice versa.

sented by the diamond). As Z0V increase, the marginal probability values for these parameters decrease rapidly, indicating the model exceeds observational uncertainties.

The marginal probabilities for RCMIN and Z0V show strong constraints, particularly for the right boundary of these parameters. Uncertainty for minimum stomatal resistance is within 6–80 s m^{-1} for Abracof, Abracop, Amazon, ARMCART, and Cabauw sites and 20–180 s m^{-1} for the Tucson site at the 95% confidence level. Uncertainty ranges are within 0–0.1 m for Z0V for four short-vegetation sites (see Fig. 6). However, as one can see the distributions for Z0V are not Gaussian and therefore the 95% confidence interval may be an overestimate of the ''most likely'' range. Vegetation albedo and vegetation cover fraction show larger uncertainty, but they are also well constrained.

At the two grassland sites, similar marginal PPDs can be found for RCMIN and Z0V while there is a shift in the peak in the marginal PPDs for vegetation albedo. At the seven sites, more spiked distributions can be found for RCMN and Z0V, implying less uncertainty in these parameters. Large uncertainties (wide PPDs) were found for ground albedo at the seven sites (Fig. 8) and for initial soil temperature at the seven sites. The approximate uniform distribution of PPDs for snow albedo, maximum leaf average index, vegetation fraction cover seasonality, and snow roughness index indicate the largest uncertainty.

It should be noted that nonzero probabilities exist, in principle outside the explored parameter region (as indicated by the nonzero probability for vegetation fraction cover at 0.7 and/or 1.0). However, we assume here that parameter values outside the explored regions are unphysical or unreasonable and therefore are assigned to zero likelihoods.

To explore the sensitivity of these results to the length of the observational data, we separate the observed data at the Loobos and Amazon sites into two parts, one for the year 1997 and the other for 1998.

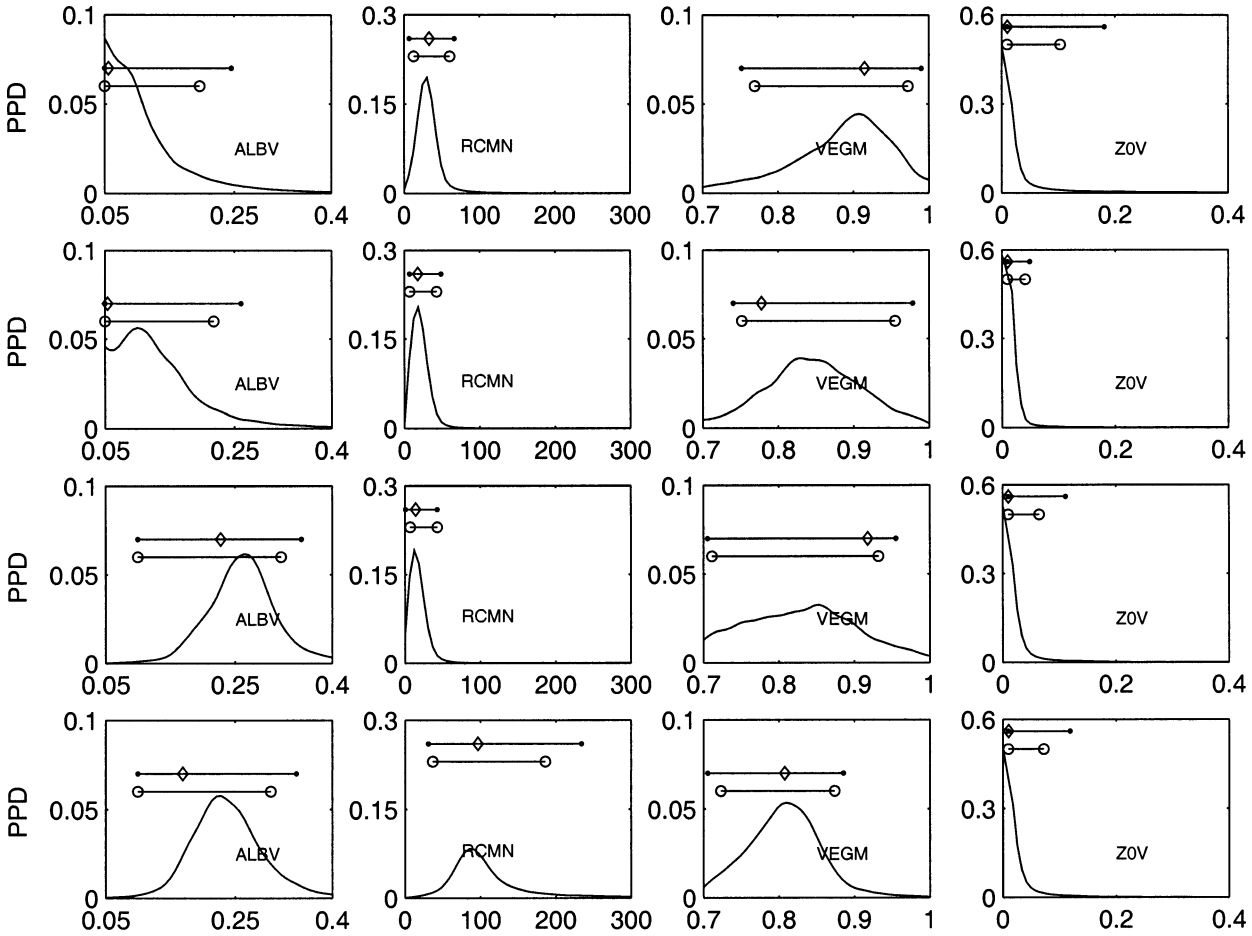

FIG. 6. Marginal PPD computed by multiple very fast simulated annealing for four parameters at three grasslands and one semiarid site. The curves were smoothed using a five-point smoothing operator (dot = 95% and circle = 90% confidence level; diamond = the best parameter set; results from stations Abracop, ARMCART, Cabauw, and Tucson are shown from the first row to the fourth row, respectively).

The calculated marginal PPDs and confidence intervals at the Loobos and Amazon sites are shown in Fig. 9. The gray line represents the year 1997 and the black line 1998. The best parameter sets are represented by the diamonds. The lines between the two dots represent the 95% confidence intervals. Although there are some differences for vegetation albedo at the Amazon site and vegetation fraction cover at the Loobos site, similar marginal PPDs and confidence intervals are obtained at the Loobos and Amazon sites for the selected four important parameters shown in Table 2 for the years 1997 and 1998. For the other parameters there are marginal differences at these two sites. This means that marginal PPDs are not significantly sensitive to the length of observational data. This is true only for the examined periods at these two sites. A proper and complete assessment on sensitivity of data length to calculation of marginal PPDs still needs a more thorough testing as done by Yapo et al. (1996) where 40-yr observational data were used.

b. Correlation matrix for the CHASM parameters

The interdependence between different model parameters can be best understood by studying the posterior correlation matrix. It is not necessary to report all the correlation coefficients for the seven sites because many of these coefficients are close to zero. Table 8 gives the correlation coefficients between those parameters whose correlation coefficients are relatively large. Vegetation albedo is negatively correlated with ground albedo (correlation coefficients from -0.2 to -0.4) and positively correlated with vegetation cover fraction (correlation coefficients from 0 up to 0.46). Because total albedo is an area-weighted average of ground and vegetation albedo, the anticorrelation and positive correlations for the respective parameters helps the model maintain the observed constraints on system energy. There also exist statistically significant correlations between the vegetation albedo and minimum stomatal resistance (correlation coefficients from 0.2 to 0.6). The reasons for this

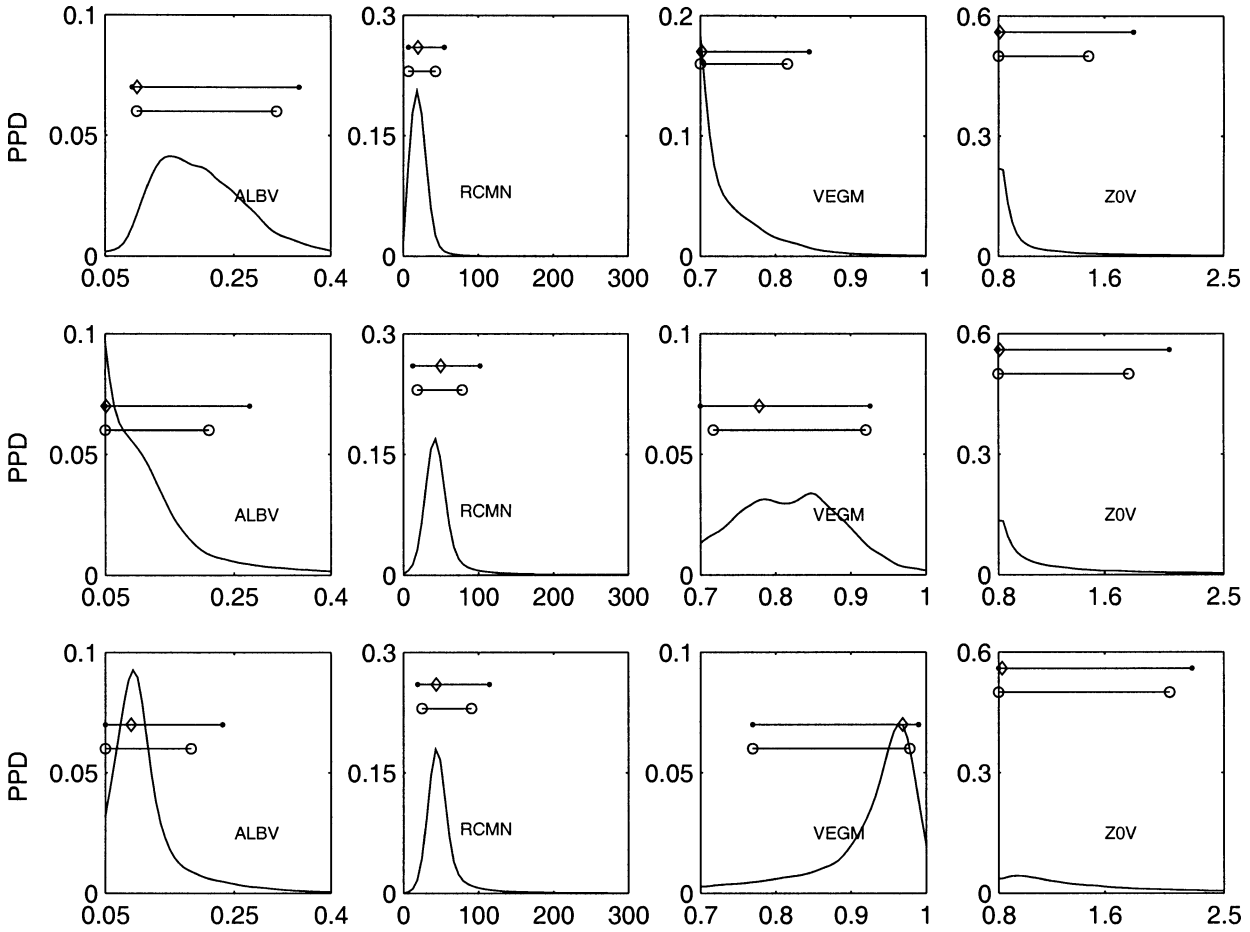

FIG. 7. Same as in Fig. 5 but for two tropical forests and one midlatitude pine forest. The curves were smoothed using a five-point smoothing operator (dot = 95% and circle = 90% confidence level; diamond = the best parameter set; results from Abracof, Amazon, and Loobos are shown from the first row to the third row, respectively).

correlation are not clear because one parameter (ALBV) affects the total energy fluxes and the other (RCMN) mainly controls the partitioning between sensible and latent heat fluxes. No other significant correlations were found between the other model parameters. Note that the correlation is strictly a linear measure of the dependence between two parameters. For parameters that are nonlinearly related, the correlation value may not indicate fully the parameters interdependence.

6. Impacts of initial soil moisture and nonmosaic approach on optimal parameters and their uncertainty analysis

Because the computational expense of parameter optimization and uncertainty estimation can be quite large, sometimes it is necessary to limit the analysis to a few key parameters. One may use different sensitivity analyses to estimate the potential importance of a given parameter, but potentially important parameters or sources of uncertainty may be unintentionally neglected. Within this discussion, we consider two potentially important factors that were initially neglected within the analyses presented earlier, initial soil moisture and the mosaic versus nonmosaic approach to constructing land surface physics parameterizations. Initial soil moisture has the potential to affect ground evaporation for many months. Moreover, it influences latent heat flux and partitioning of surface energy fluxes. Niyogi et al. (2002) discussed the importance of initial soil moisture on sensible and latent heat fluxes with the Simplified Simple Biosphere model and its relationship with vegetation fraction and minimum stomatal resistance in a midlatitude regime. In order to examine impact of initial soil moisture on the optimization processes and uncertainty estimates of the CHASM, we conducted a second set of experiments that includes initial soil moisture within the BSI optimization and uncertainty estimation for all sites except Abracof. (Because Abracof and Amazon sites are in the Amazonian tropical forests, they have similar climate, vegetation, and soil types. Therefore, results at Amazon are representative for the Abracof site.) The additional experiments confirm that initial soil moisture is an important parameter for the sites consid-

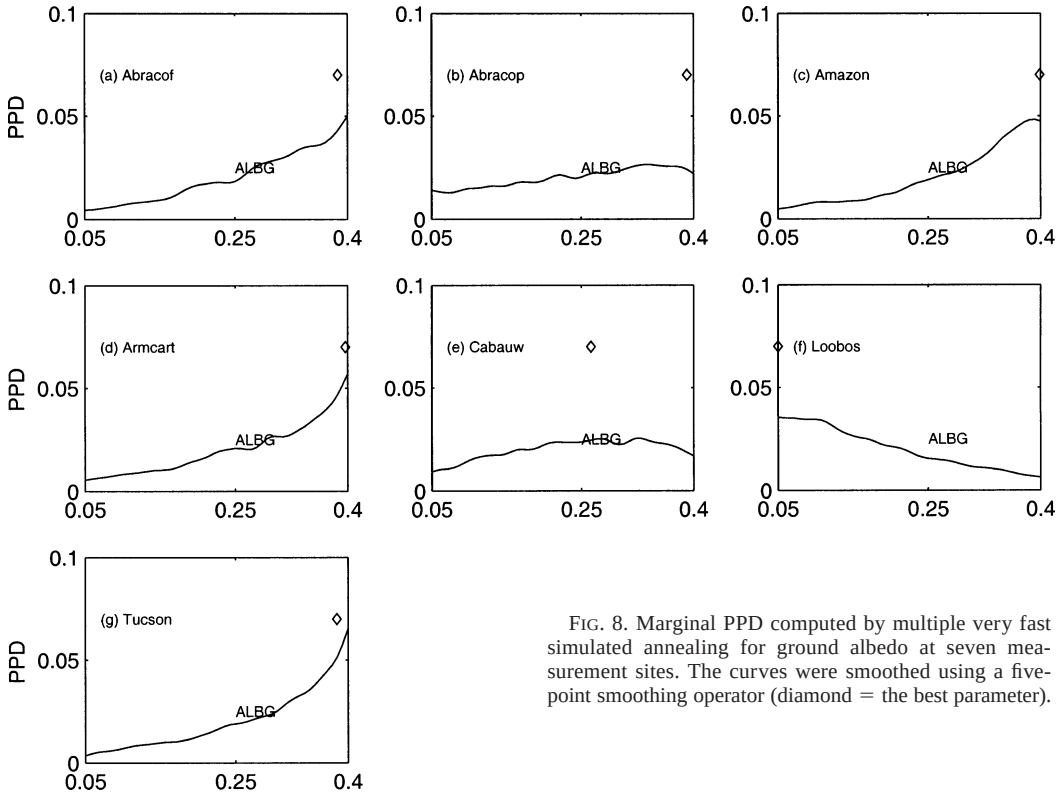

FIG. 8. Marginal PPD computed by multiple very fast simulated annealing for ground albedo at seven measurement sites. The curves were smoothed using a five-point smoothing operator (diamond = the best parameter).

ered, consistent with results of Niyogi et al. (2002) for midlatitude and arid tropical regimes. Initial soil moisture was only found to have a statistically significant correlation with maximum water-holding capacity. Therefore, while uncertainty in initial soil moisture can contribute to predictive uncertainty of the CHASM model, it does not strongly affect the choice of other parameter values. In this study, we do not find significant correlations between initial soil moisture and other parameters such as vegetation cover fraction, minimum stomatal resistance, ground albedo, and soil temperature as reported by Niyogi et al. (2002). The reason for the inability to find these relationships is not clear and it needs to be further investigated using different land surface models and different sensitivity analysis methods in the future.

Another potentially important difference between different formulations of land surface models is the adoption of the so-called mosaic approach to physical parameterization. Both nonmosaic and mosaic approaches (Koster and Suarez 1992) were considered in the PILPS offline set of experiments (Chen et al. 1997; Wood et al. 1998; Schlosser et al. 2000; Slater et al. 2001; Bowling et al. 2003). The nonmosaic approach defines a set of effective surface parameters (e.g., roughness length, albedo, surface resistance to evaporation) that provide a close approximation to the grid box mean fluxes from patchy landscapes. Effective parameters are calculated offline as the area-weighted mean values of the surface types in a grid box for the CHASM. For some parameters (e.g., roughness lengths), more sophisticated methods are used because of the nonlinear relationship between the fluxes and the parameters. An energy balance equation is solved to calculate fluxes as grid box means. In contrast, the mosaic approach solves separate surface energy balances for each surface type (or tile) and area-weighted means of the fluxes from tiles are used to represent grid box mean values. So the mosaic approach aggregates fluxes, whereas the nonmosaic approach aggregates parameters.

The net effect of uncertainties in initial soil moisture and the mosaic versus nonmosaic approaches on rmse and biases for sensible and latent heat fluxes at six measurement sites are shown in Fig. 10. The case with the nonmosaic approach includes the optimization for initial soil moisture. The results show that rmse of sensible and latent heat fluxes are similar for the three experiments. However, there are significant differences among the three experiments for biases of sensible and latent heat fluxes at some sites. The latent heat fluxes at ARMCART and Cabauw are two good examples: the negative/positive biases (at ARMCART/Cabauw) are reduced for the varied soil moisture case and are increased again for the nonmosaic approach case. However, the biases of sensible fluxes are reduced for the nonmosaic approach case from varied soil moisture cases at the two sites. Overall, use of varied soil moisture leads to equal

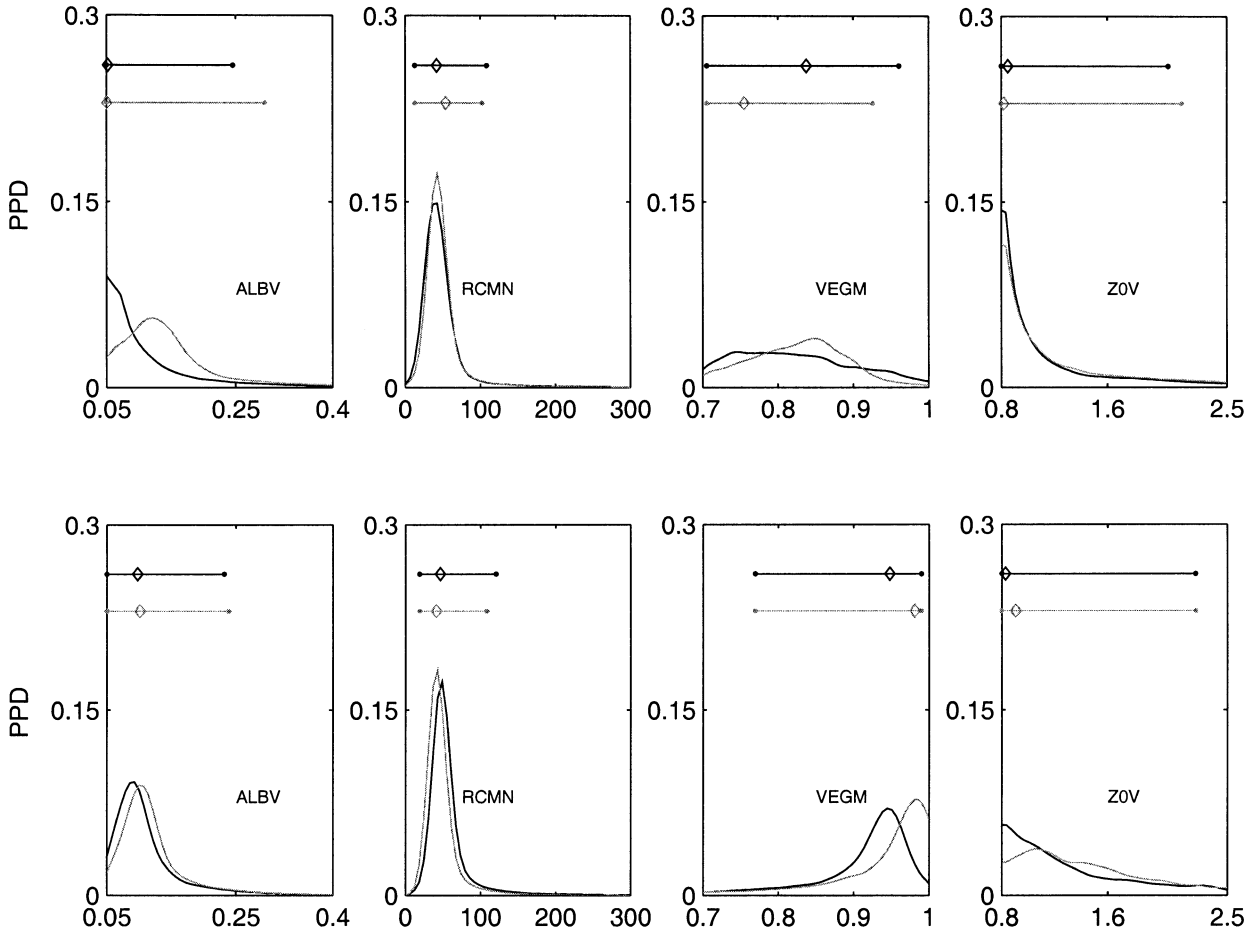

FIG. 9. Same as in Fig. 5 but for Amazon tropical forests and Loobos midlatitude pine forest. (dot = 95% confidence level, diamond = the best parameter set, black = the year 1997, and gray = 1998; results from Amazon and Loobos are shown from the first and the second row, respectively).

or better simulations when compared with results where soil moisture is fixed.

Figure 11 considers the impacts of these uncertainties on the marginal PPDs derived for minimum stomatal resistance (RCMN). RCMN is an important parameter for all sites insofar as it impacts the partitioning of sensible and latent heat fluxes at the land surface. In Fig. 11 the solid lines represent the case of fixed initial soil moisture, solid-dotted lines represent the case of varied initial soil moisture, and gray lines represent the case of the nonmosaic model. The results show that use of varied initial soil moisture does not change shapes of the marginal PPDs for any of the examined sites except for Tucson. Instead, the principal change is in the width of the distributions. This means that including initial soil moisture within the parameter optimization increases the uncertainty ranges derived for minimum stomatal resistance. Because there was not much change between the distributions between the varied initial soil moisture and the nonmosaic approach (except again for Tucson), we can infer that this structural change in the land surface formulation is not critical. At most of the sites considered here besides Tucson, vegetation covers over 90% of land surface. Therefore, one may expect similar marginal PPDs of minimum stomatal resistance between the mosaic and nonmosaic approaches. At Tucson, the nonmosaic approach significantly changes the shape of the PPD and uncertainty ranges of minimum stomatal resistance when compared with mosaic results (Fig. 11f). The point of the largest likelihood is shifted from

TABLE 8. Correlation coefficients between vegetation albedo (ALBV) and other selected parameters for seven sites.

Sites	Abracof	Abracop	Amazon	ARMCART	Cabauw	Loobos	Tucson
ALBG	−0.36	−0.27	−0.27	−0.20	−0.40	−0.18	−0.41
VEGM	0.46	0.33	0.20	0.23	0.15	0.01	0.30
RCMN	0.53	0.36	0.28	0.59	0.39	0.25	−0.19

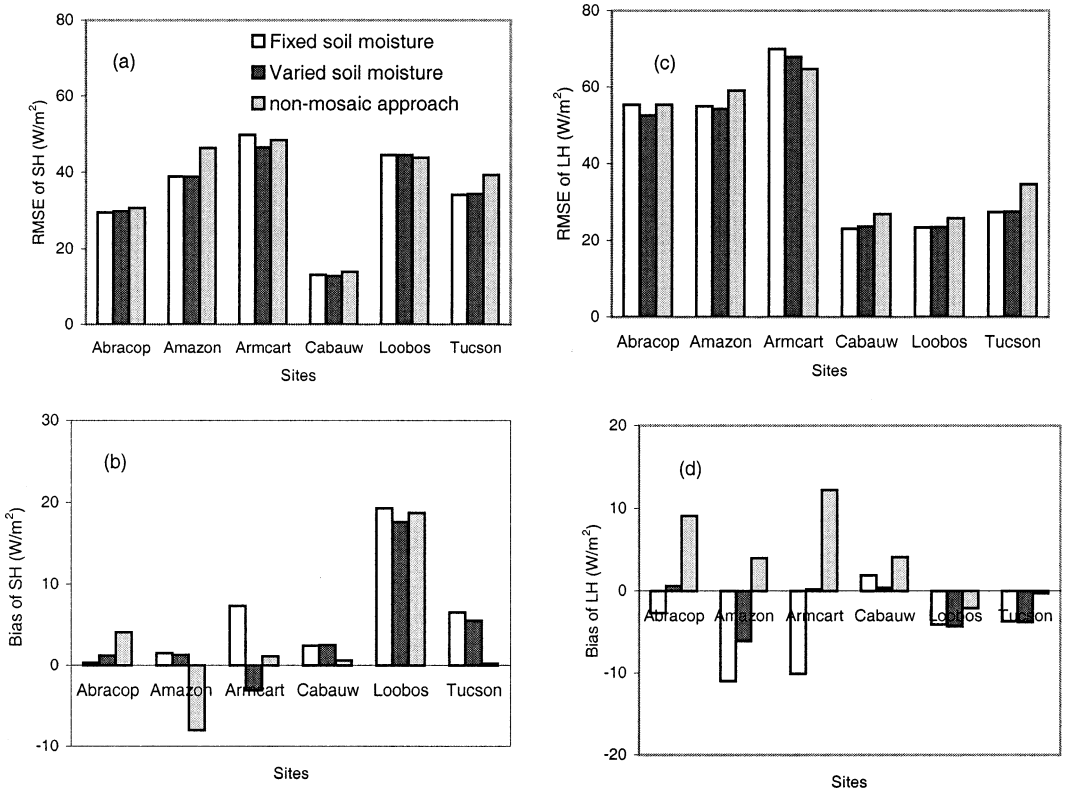

FIG. 10. Rmse for (a) SH and (c) LH fluxes, and biases for (b) SH and (d) LH fluxes calculated by optimal parameter sets calibrated with fixed and varied initial soil moisture, and nonmosaic approach for the six examined sites.

100 to 20 s m^{-1}. The importance of the mosaic versus nonmosaic approaches on the PPD is physically reasonable at this location because vegetation covers about 60% of land surface with bare ground covering about 40% of land surface. Therefore, one may more generally expect that the average of parameters (nonmosaic) and average of energy fluxes (mosaic) between separately defined patches should result in different distributions of marginal PPD's distributions of minimum stomatal resistance.

7. Discussion

The uncertainty ranges of sensible and latent heat fluxes in our study are comparable to PILPS model scatter (ranges of all 23 land surface models; Chen et al. 1997; Desborough 1999) for monthly and annual mean results at the Cabauw site. Uncertainty sources for both PILPS scatter and uncertainty ranges of our study may come from the combination of uncertainty in values assigned to important parameters, model structure differences, uncertainty within observed data, and so on, because these uncertainty sources cannot be easily separated.

This offline study is similar to Frank and Beven (1997) but is more comprehensive in that it considers more biomes and potential sources of uncertainty. On average we tested ~100 000 model parameter sets chosen stochastically by a VFSA algorithm that permits one to concentrate on regions of the estimated PPD that are the most important for each site and each experiment. In contrast, Franks and Beven (1997) drew 10 000 parameter sets randomly (Monte Carlo sampling) for 15 parameters of a land surface model. The purely random Monte Carlo sampling strategy has been found to be 2–3 orders of magnitude less efficient in comparison with the type of importance sampling that can be obtained with VFSA (see Sen and Soffa 1995, 1996; Jackson et al. 2003, 2004).

Second, our analysis considers long-term observational data covering 6 months–2 yr at seven sites, which are representative of tropical forests, tropical pastures, midlatitude grasslands, midlatitude crops, semiarid grasslands, and midlatitude pine forests, while Franks and Beven (1997) consider shorter-term observational data with 2 weeks at tropical forest and midlatitude grassland sites that are different from our sites. Short-term field campaigns represented by these datasets may not be adequate to specify parameter values characteristic of a site or area as reported by Franks and Beven (1997). Therefore, longer periods of data as we used here, incorporating obvious seasonal variation of calibration data, may be more appropriate to constrain the feasible parameter sets or model structures.

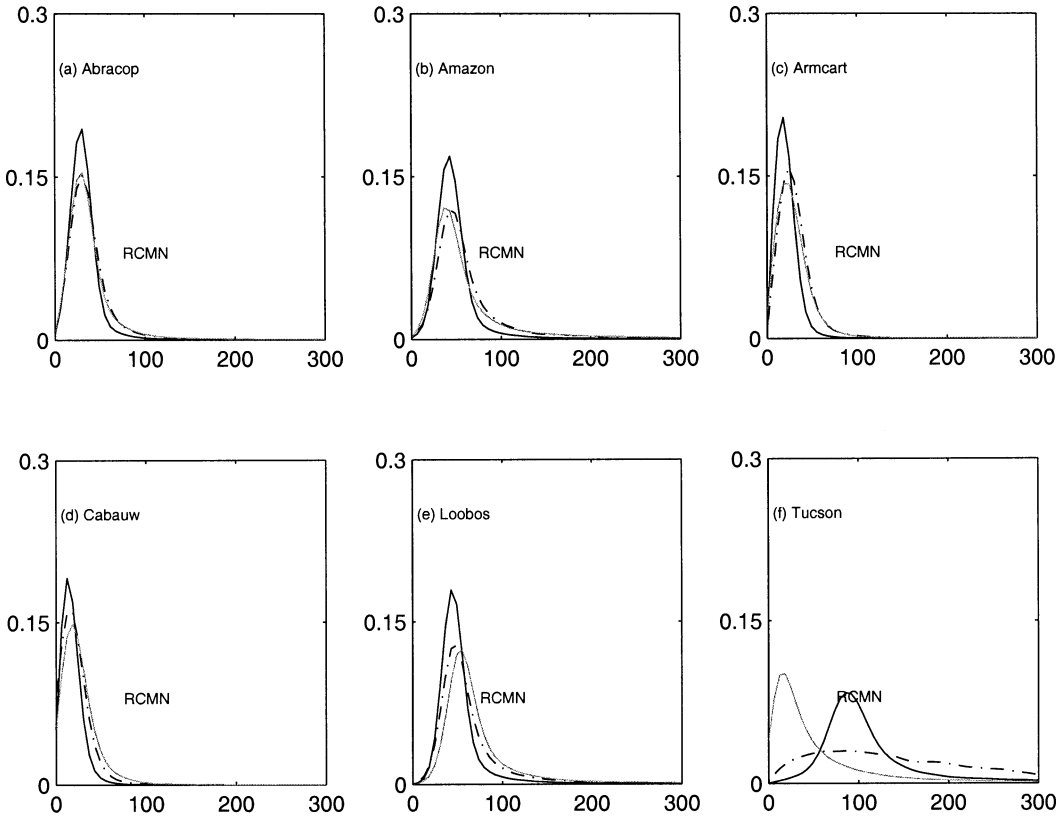

FIG. 11. Calculated marginal PPDs of minimum stomatal resistance for fixed and varied initial soil moisture and nonmosaic approach at (a) Abracop, (b) Amazon, (c) ARMCART, (d) Cabauw, (e) Loobos, and (f) Tucson sites (solid line = fixed initial soil moisture, dashed–dotted line = varied initial soil moisture, and gray line = nonmosaic approach).

The estimation of the potential sources of parameter uncertainties and their impacts on surface energy fluxes that we document for the CHASM are conditional on several subjective elements. The results apply exclusively to the CHASM and reflect the period and quality of the observational datasets as well as the biases inherent in using VFSA to select parameter values. In addition, like other approaches to quantifying uncertainty using Bayesian statistics (i.e., GLUE), BSI contains a number of subjective choices that may, to a limited extent, affect the analysis. These more subjective elements include the choice of parameter ranges to consider and the choice of cost (error) function definition. Therefore, the parameter values and ranges reported in this paper should be used in this context.

8. Conclusions

We investigate the ability of BSI and MC to identify the optimal parameter sets for the CHASM at seven sites. Our results indicate that estimated optimal values are very similar for the important parameters although parameters values may be significantly different for insensitive parameters. Calibrations improve the simulations of sensible and latent heat fluxes for the CHASM regardless of the calibration method at five of the seven sites. Cross-validation of two tropical forest sites suggests that the calibrated parameters can be transferred to alternate sites within the same biomes. This spatial transferability justifies, in part, the use of field-calibrated parameters within GCM simulations (e.g., Sen et al. 2001).

Bayesian stochastic inversion is used to calculate marginal PPDs for CHASM parameters and correlation matrices between model parameters. Marginal PPDs are used to estimate and analyze the source of uncertainties arising from arbitrary choices of parameters of the CHASM, and correlation matrices are used to investigate the relationships between the CHASM parameters. The results demonstrate that the CHASM shows significant uncertainty for important model parameters, such as minimum stomatal resistance (6–120 s m^{-1}), vegetation roughness length (0.0–0.01 for grasslands), and ground roughness length (0.0–0.08 m). Analysis of correlation matrices shows a negative correlation between vegetation albedo and ground albedo, a positive correlation between the vegetation albedo and minimum stomatal resistance, and a positive correlation between vegetation albedo and vegetation cover fraction for almost all sites. Examinations of multiyear datasets indicate that the use of 1-yr datasets may be appropriate for estimating the distributions of marginal PPDs of

CHASM parameters. We used the parameter sets selected by the BSI to estimate uncertainty ranges of seasonally averaged, monthly mean, and annual mean sensible and latent heat fluxes at seven sites at the 95% confidence level and discuss the effect of initial soil moisture and the nonmosaic approach on optimization processes and uncertainty estimates of the CHASM.

Acknowledgments. Authors YX and CJ were supported by G. Unger Vetlesen Foundation. Special thanks go to Dr. Hoshin Gupta for providing the multicriteria methodology, Dr. Andy Pitman for providing the CHASM, and Dr. Luis Bastidas for providing ARM-CART and Tucson datasets. The Royal Netherlands Meteorological Institute kindly provided the Cabauw dataset. Abracof and Abracop data were collected under the ABRACOS project and made available by the U.K. Institute of Hydrology and the Instituto Nacional de Pesquisas Espaciais (Brazil). ABRACOS is a collaboration between the Agencia Brasileira de Cooperacao and the UK Overseas Development Administration. We thank three anonymous referees whose comments greatly improved the quality of this paper.

APPENDIX

Basic Parameterization Related to the Latent Heat Flux

a. Basic parameterization

CHASM uses a grouped mosaic approach (Koster and Suarez 1992) to resolve the surface energy balance to a soil depth of 10 cm. Similar elements from a grid cell are put together to form tiles (e.g., vegetation, bare soil), and a separate surface energy balance equation is developed and resolved for each tile. Each tile has a prognostic bulk temperature for storage of energy and a diagnostic skin temperature for calculation of surface energy fluxes. Each tile can have up to four evaporation sources related to latent heat fluxes. Transpiration (E_{tr}), canopy evaporation (E_c), bare-ground evaporation (E_g), and snow sublimation (E_n) are expressed as

$$E_{\rm tr}^{[t]} = \frac{A^{[t]}(1 - a_{\rm wet}^{[t]})a_v^{[t]}\rho_a\beta_{\rm tr}(q^{*[t]} - q_a)}{(r_a + r_c^{*[t]})}, \quad \text{(A1)}$$

$$E_c^{[t]} = \frac{A^{[t]}a_{\rm wet}^{[t]}a_v^{[t]}\rho_a(q^{*[t]} - q_a)}{r_a}, \quad \text{(A2)}$$

$$E_g^{[t]} = \frac{A^{[t]}a_g^{[t]}\rho_a\beta_g^{[t]}(q^{*[t]} - q_a)}{r_a}, \quad \text{and} \quad \text{(A3)}$$

$$E_n^{[t]} = \frac{A^{[t]}a_n^{[t]}\rho_a(q^{*[t]} - q_a)}{r_a}, \quad \text{(A4)}$$

where $A^{[t]}$ is a tile fraction. Each tile fraction is further divided into vegetation ($a_v^{[t]}$), ground ($a_g^{[t]}$), and snow ($a_n^{[t]}$) fractions. If canopy interception occurs, vegetation fraction is further divided into wet ($a_{\rm wet}^{[t]}$) and dry ($1 - a_{\rm wet}^{[t]}$) fractions each time step and $a_{\rm wet}^{[t]}$ is set equal to the canopy interception store's fractional wetness. The values $E_{\rm tr}^{[t]}$ and $E_g^{[t]}$ are reduced below their potential rates by moisture availability indices ($\beta_{\rm tr}$ and $\beta_g^{[t]}$), and $E_{\rm tr}^{[t]}$ must overcome an additional resistance ($r_c^{*[t]}$) that represents the influence of nonmoisture stomatal stresses. Here, r_a is an aerodynamic resistance for heat and moisture. The surface saturated specific humidity ($q^{*[t]}$) is obtained as a function of diagnostic skin temperature. The values ρ_a and q_a are air density and air specific humidity of forcing data, respectively.

b. Canopy resistance

Following the basic philosophy of Jarvis (1976), $r_c^{*[t]}$ is expressed as the product of a minimum value ($r_{c\min}$) and a series of environmental factors for radiation (f_R), temperature (f_T), and humidity (f_H):

$$r_c^{*[t]} = r_{c\min}/(f_R f_T f_H), \quad \text{(A5)}$$

$$f_R = F_s/200, \quad \text{(A6)}$$

$$f_T = 1 - (298 - T_3)^2, \quad \text{and} \quad \text{(A7)}$$

$$f_H = 500/V_D, \quad \text{(A8)}$$

where each of the functions is constrained to lie between 0 and 1 and $r_c^{*[t]}$ is constrained to lie between $r_{c\min}$ and a maximum value of 1000 s m^{-1}. Here F_s is incident solar radiation, T_3 is soil temperature at 1 m, and V_D is the near-surface air vapor pressure deficit. The $r_{c\min}$ is a sensitive and important parameter in optimization and uncertainty analysis processes of our study.

c. Aerodynamic resistance to turbulent transport

The aerodynamic resistance (r_a) for heat and moisture is parameterized as

$$r_a = 1/(a^2 F_h u_a), \quad \text{with} \quad \text{(A9)}$$

$$a^2 = [k_{\rm vk}/\log(z_a + z_0)/z_0]^2, \quad \text{(A10)}$$

where z_a is the atmospheric forcing height, z_0 is the surface roughness length and $k_{\rm vk}$ is von Kármán's constant. The value F_h is used to parameterize the dependence of r_a on atmospheric stability, with the resistance to turbulent exchange decreasing as thermal forcing increases.

$$F_h = 1/[1 + 10R_{\rm iB}(1 + 8R_{\rm iB})] \quad \text{when } T_s < T_a, \quad \text{(A11)}$$

$$F_h = 1 - 15R_{\rm iB}/(1 + c|R_{\rm iB}|^{1/2}) \quad \text{when } T_s > T_a, \quad \text{(A12)}$$

and (A13)

$$F_h = 1 \quad \text{when } T_s = T_a,$$

where T_s and T_a are surface temperature and air temperature, respectively. The bulk Richardson number ($R_{\rm iB}$) and adjustable parameter (c) are expressed as

$$R_{iB} = gz_a(T_a - T_s)/(T_s u_a^2) \quad \text{and} \qquad \text{(A14)}$$

$$c = 75a^2[(z_a + z_0)/z_0]^{1/2}, \qquad \text{(A15)}$$

where g is acceleration due to gravity. Roughness length (z_0) is calculated as log-weighted average across vegetation, bare ground, and snow fractions. Roughness lengths for vegetation, bare soil, and snow are optimized using BSI and MC approaches. Vegetation roughness length is an important parameter for optimization and uncertainty analysis processes of our study.

REFERENCES

Alapaty, K., S. Raman, and D. S. Niyogi, 1997: Uncertainty in specification of surface characteristics: A study of prediction errors in the boundary layer. *Bound.-Layer Meteor.,* **82,** 473–500.

Bastidas, L. A., H. V. Gupta, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang, 1996: The land surface–atmosphere interaction: A review based on observational and global modeling perspectives. *J. Geophys. Res.,* **101,** 7209–7225.

——, ——, ——, ——, and ——, 1999: Sensitivity analysis of a land surface scheme using multi-criteria methods. *J. Geophys. Res.,* **104,** 19 481–19 490.

Bates, B. C., and E. P. Campbell, 2001: A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. *Water Resour. Res.,* **37,** 937–947.

Beljaars, A. C. M., and F. Bosveld, 1997: Cabauw data for the validation of land surface parameterization schemes. *J. Climate,* **10,** 1172–1193.

Beringer, J., S. McIlwaine, A. H. Lynch, F. S. Chapin III, and G. B. Bonan, 2002: The use of a reduced form model to assess the sensitivity of a land surface model to biotic surface parameters. *Climate Dyn.,* **19,** 455–466.

Bonan, G. B., D. Pollard, and S. L. Thompson, 1993: Influence of subgrid-scale heterogeneity in leaf area index, stomatal resistance, and soil moisture on grid-scale land–atmosphere interactions. *J. Climate,* **6,** 1882–1897.

Bowling, L. C., and Coauthors, 2003: Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e) 1. Experiment description and summary intercomparisons. *Global Planet. Change,* **38,** 1–35.

Campbell, E. P., D. R. Fox, and B. C. Bates, 1999: A Bayesian approach to parameter estimation and pooling in nonlinear flood event models. *Water Resour. Res.,* **35,** 211–220.

Chen, T. H., and Coauthors, 1997: Cabauw experimental results from the Project for Intercomparison of Land-Surface Parameterization Schemes. *J. Climate,* **10,** 1194–1215.

Collins, D. C., and R. Avissar, 1994: An evaluation with the Fourier amplitude sensitivity test (FAST) of which land-surface parameters are of greatest importance in atmospheric modeling. *J. Climate,* **7,** 681–703.

Crossley, J. F., J. Polcher, P. M. Cox, N. Gedney, and S. Planton, 2000: Uncertainties linked to land-surface processes in climate change simulations. *Climate Dyn.,* **16,** 949–961.

Desborough, C. E., 1999: Surface energy balance complexity in GCM land surface models. *Climate Dyn.,* **15,** 389–403.

——, A. J. Pitman, and B. McAvaney, 2001: Surface energy balance complexity in GCM land surface models. Part II: Coupled simulations. *Climate Dyn.,* **17,** 615–626.

Duan, Q., V. K. Gupta, and S. Sorooshian, 1992: Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.,* **28,** 1015–1031.

——, S. Sorooshian, and V. K. Gupta, 1994: Optimal use of the SCE-UA global optimization method for calibrating watershed models. *J. Hydrol.,* **158,** 265–284.

Franks, S. W., and K. J. Beven, 1997: Bayesian estimation of uncertainty in land surface–atmosphere flux predictions. *J. Geophys. Res.,* **102,** 23 991–23 999.

Gao, X., S. Sorooshian, and H. V. Gupta, 1996: Sensitivity analysis of the Biosphere–Atmosphere Transfer Scheme. *J. Geophys. Res.,* **101,** 7279–7289.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2003: *Bayesian Data Analysis.* 2d ed. Chapman and Hall, 696 pp.

Gupta, H. V., S. Sorooshian, and P. O. Yapo, 1998: Toward improved calibration of hydrologic models: Multiple and non-commensurable measures of information. *Water Resour. Res.,* **34,** 751–761.

——, L. A. Bastidas, S. Sorooshian, W. J. Shuttleworth, and Z.-L. Yang, 1999: Parameter estimation of a land surface scheme using multicriteria methods. *J. Geophys. Res.,* **104,** 19 491–19 503.

Hastings, W. K., 1970: Monte Carlo sampling method using Markov chains and their applications. *Biometrika,* **57,** 97–109.

Henderson-Sellers, A., 1993: A factorial assessment of the sensitivity of the BATS land-surface parameterization scheme. *J. Climate,* **6,** 227–247.

——, 1996: Soil moisture simulation: Achievements of the RICE and PILPS intercomparison workshop and future directions. *Global Planet. Change,* **13,** 99–115.

——, K. McGuffie, and A. Pitman, 1996: The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS): 1992–1995. *Climate Dyn.,* **12,** 849–859.

Houghton, J. T., L. G. Meria Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., 1996: *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, 572 pp.

Jackson, C., Y. Xia, M. K. Sen, and P. Stoffa, 2003: Optimal parameter estimation and uncertainty analysis of a land surface model: A case study from Cabauw, Netherlands. *J. Geophys. Res.,* **108,** 4583, doi:10.1029/2002JD002991.

——, M. Sen, and P. Stoffa, 2004: An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions. *J. Climate,* **17,** 2828–2841.

Jarvis, P. G., 1976: The interpretation of the variations in the leaf water potential and stomatal conductance found in canopies in the field. *Philos. Trans. Roy. Soc. London,* **273B,** 593–610.

Koster, R. D., and M. J. Suarez, 1992: Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *J. Geophys. Res.,* **97,** 2697–2715.

Kuczera, G., and E. Parent, 1998: Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm. *J. Hydrol.,* **211,** 69–85.

Leplastrier, M., A. J. Pitman, H. V. Gupta, and Y. Xia, 2002: Exploring the relationship between complexity and performance in a land surface model using the multicriteria method. *J. Geophys. Res.,* **107,** 4443, doi:10.1029/2001JD000931.

Manabe, S., 1969: Climate and the ocean circulation. I: The atmospheric circulation and the hydrology of the earth's surface. *Mon. Wea. Rev.,* **97,** 739–805.

Margulis, S. A., and D. Entekhabi, 2001: Feedback between the land-surface energy balance and atmospheric boundary layer diagnosed through a model and its adjoint. *J. Hydrometeor.,* **2,** 599–619.

McWilliam, A. L. C., O. M. R. Cabral, B. M. Gomes, J. L. Esteves, and J. M. Roberts, 1996: Forest and pature leaf–gas exchange in south-west Amazonia. *Amazon Deforestation and Climate,* J. H. C. Gash et al., Eds., John Wiley, 265–286.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, 1953: Equation of state calculations by fast computing machines. *J. Chem. Phys.,* **21,** 1087–1092.

Niyogi, D. S., S. Raman, and K. Alapaty, 1998: Comparison of four different stomatal resistance schemes using FIFE data. Part II: Analysis of terrestrial biospheric–atmospheric interactions. *J. Appl. Meteor.,* **37,** 1301–1320.

——, ——, and ——, 1999: Uncertainty in specification of surface

characteristics, Part 2: Hierarchy of interaction explicit statistical analysis. *Bound.-Layer Meteor.,* **91,** 341–366.

——, Y. K. Xue, and S. Raman, 2002: Hydrological land surface response in a tropical regime and a midlatitudinal regime. *J. Hydrometeor.,* **3,** 39–56.

Pitman, A. J., 1994: Assessing the sensitivity of a land-surface scheme to the parameter values using a single column model. *J. Climate,* **7,** 1856–1869.

——, Y. Xia, M. Leplastrier, and A. Henderson-Sellers, 2003: The CHAmeleon Surface Model (CHASM): Description and use with the PILPS Phase 2e forcing data. *Global Planet. Change,* **38,** 121–135.

Roberts, J. M., O. M. R. Cabral, and L. F. de Aguiar, 1990: Stomatal and boundary layer conductances measured in a terra firme rain forest. *J. Appl. Ecol.,* **27,** 336–353.

——, ——, J. P. da Costa, A. C. L. McWilliam, and T. D. A. Sa, 1996: An overview of the leaf area index and physiological measurements during ABRACOS. *Amazon Deforestation and Climate,* J. H. C. Gash et al., Eds., John Wiley and Sons, 287–306.

Schlosser, C. A., and Coauthors, 2000: Simulations of a boreal grassland hydrology at Valdai, Russia: PILPS phase 2(d). *Mon. Wea. Rev.,* **128,** 301–321.

Sellers, P. J., and Coauthors, 1997: Modeling the exchanges of energy, water and carbon between continents and the atmosphere. *Science,* **275,** 502–509.

Sen, M. K., and P. L. Stoffa, 1995: *Global Optimization Methods in Geophysical Inversion.* Elsevier, 281 pp.

——, and ——, 1996: Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion. *Geophys. Prospect.,* **44,** 313–350.

Sen, O. L., L. A. Bastidas, W. J. Shuttleworth, Z.-L. Yang, H. V. Gupta, and S. Sorooshian, 2001: Impact of field-calibrated vegetation parameters on GCM climate simulations. *Quart. J. Roy. Meteor. Soc.,* **127,** 1199–1223.

Shuttleworth, W. J., 1984: Observations of radiation exchange above and below Amazonian forest. *Quart. J. Roy. Meteor. Soc.,* **110,** 1163–1169.

Skaggs, T. H., and D. A. Barry, 1996: Sensitivity methods for time-continuous, spatially discrete groundwater contaminant transport models. *Water Resour. Res.,* **32,** 2409–2420.

Slater, A. G., and Coauthors, 2001: The representation of snow in land surface schemes: Results from PILPS 2(d). *J. Hydrometeor.,* **2,** 7–25.

Thiemann, M., M. Trosser, H. Gupta, and S. Sorooshian, 2001: Bayesian recursive parameter estimation for hydrologic models. *Water Resour. Res.,* **37,** 2521–2535.

Unland, H. E., P. R. Houser, W. J. Shuttleworth, and Z. L. Yang, 1996: Surface flux measurement and modeling at a semi-arid Sonoran Desert site. *Agric. For. Meteor.,* **82,** 119–153.

Wang, Q. J., 1991: The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resour. Res.,* **27,** 2467–2471.

Wilson, M. F., A. Henderson-Sellers, R. E. Dickinson, and P. J. Kennedy, 1987: Sensitivity of the Biosphere–Atmosphere Transfer Scheme (BATS) to the inclusion of variable soil characteristics. *J. Climate Appl. Meteor.,* **26,** 341–362.

Wood, E. F., and Coauthors, 1998: The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS) phase 2(c) Red–Arkansas River basin experiment. I. Experiment description and summary intercomparisons. *Global Planet. Change,* **19,** 115–135.

Xia, Y., A. J. Pitman, H. V. Gupta, M. Leplastrier, A. Henderson-Sellers, and L. A. Bastidas, 2002: Calibrating a land surface model of varying complexity using multicriteria methods and the Cabauw dataset. *J. Hydrometeor.,* **3,** 181–194.

Yapo, P. O., H. V. Gupta, and S. Sorooshian, 1998: Multi-objective global optimization for hydrologic models. *J. Hydrol.,* **204,** 83–97.

Zhang, H., A. Henderson-Sellers, A. J. Pitman, J. L. McGregor, C. E. Desborough, and J. J. Katzfey, 2001: Limited-area model sensitivity to the complexity of representation of the land surface energy balance. *J. Climate,* **14,** 3965–3986.